## (19) World Intellectual Property Organization

International Bureau





(43) International Publication Date 21 May 2004 (21.05.2004)

PCT

(10) International Publication Number WO 2004/042365 A2

(51) International Patent Classification<sup>7</sup>:

**G01N** 

(21) International Application Number:

PCT/US2003/036247

(22) International Filing Date:

3 November 2003 (03.11.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data: 60/423,113 1 Nov

1 November 2002 (01.11.2002) US

(71) Applicant (for all designated States except US): EVOLUTIONARY GENOMICS LLC [US/US]; Bioscience Park Center, 12635 East Montview Boulevard, Suite 211, Aurora, CO 80010 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): MESSIER, Walter [US/US]; 2418 Bowen Street, Longmont, CO 80501 (US).

(74) Agents: SWANSON, Barry, J. et al.; Swanson & Bratschun, L.L.C., 1745 Shea Center Drive, Suite 330, Highlands Ranch, CO 80129 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DEVELOPMENT OF THERAPEUTICS FOR THE TREATMENT OF ENDOTOXIN-MEDIATED DISEASES

(57) Abstract: The subject invention comprises a method for identifying an evolutionarily meaningful nucleotide change in a primate's TLR4 polynucleotide. It further comprises methods for identifying agents that interact with the corresponding evolutionarily meaningful amino acid change so as to modulate the function of the TLR4 polypeptide, thereby attenuating activation of the NF-kB pathway. Such agents are useful in mitigating the LPS mediated response and in the treatment of sepsis, severe sepsis and septic shock.

**NO 2004/042365** A:

## DEVELOPMENT OF THERAPEUTICS FOR THE TREATMENT OF ENDOTOXIN-MEDIATED DISEASES

## FIELD OF THE INVENTION

The present invention relates to methods to develop agents for treating endotoxin-mediated diseases of humans, such that the therapeutic agent interacts with the human TLR4 polypeptide extracellular domain in such a way as to attain a response resembling that of the Old World monkeys, specifically baboons and rhesus monkeys.

10

15

5

## BACKGROUND OF THE INVENTION

Sepsis is a serious medical condition. According to R.A. Balk and L. C. Casey (April 2000 Critical Care Clinics):

- Sepsis results in 120,000 to 200,000 deaths annually in the United States
- Death due to this discase (4.2 deaths/100,000) has increased 82.6% from 1979 to 1997.
- It is the 12th leading cause of death overall and is the most common cause of shock encountered by internists in the U.S.
- Between 300,000 to 500,000 cases of sepsis are diagnosed per year
- Shock develops in about 40% of septic patients.
- Despite aggressive treatment, mortality ranges from 16% in patients with sepsis to 40-60% in patients with septic shock
- The annual health care cost from caring for patients with sepsis is \$5-10 billion

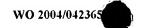
25

20

To date, no truly effective therapy exists to counteract the effects of sepsis, although some techniques do show limited utility. An effective therapeutic approach would have tremendous social and commercial value. Described here is a method to develop to such a therapeutic.

30

Severe sepsis, also known as septic syndrome, refers to a chain of events leading from microbial infection to tissue injury and cardiovascular collapse. J.S. Stapczynski provides the following definitions:



10

15

20

25

30



- "Sepsis" is the systemic host response to infection, defined as SIRS (systemic inflammatory response syndrome) in combination with a documented infection
- "Severe sepsis" is defined as sepsis plus end-organ dysfunction or hypoperfusion
- "Septic shock" is defined as sepsis with hypotension, despite fluid resuscitation, and evidence of inadequate tissue perfusion

Significant complications from sepsis include central nervous system dysfunction, adult respiratory distress syndrome (ARDS), liver failure, acute renal failure (ARF), and disseminated intravascular coagulation (DIC). In different studies, the reported incidence rates of these complications in SIRS and sepsis is about 19% for CNS dysfunction, 2-8% for ARDS, 12% for liver failure, 9-23% for ARF, and 8-18% for DIC (Stapczynski, J.S. 2001 *eMedicine Journal* 2:7).

According to N.R. Chamberlain (2001 Bacterial sepsis with shock in *Infectious Diseases Lectures*), sepsis involves a very complex sequence of events and much remains incompletely understood about how a patient goes from SIRS to septic shock. Patients with septic shock have a biphasic immunological response. Initially they manifest an overwhelming inflammatory response to an infection. This is probably due to the pro-inflammatory cytokines tumor necrosis factor (TNF), IL-1, IL-12, Interferon gamma (IFN $\gamma$ ), and IL-6). The body then regulates this response by producing anti-inflammatory cytokines (IL-10), soluble inhibitors (TNF receptors, IL-1 receptor type II, and IL-1RA, an inactive form of IL-1). The patient manifests a period of immunodepression. Persistence of this hyporesponsiveness is associated with increased risk of nosocomial infection and death.

Approximately one half of septic shock cases are caused by Gram-negative bacteria (Balk, R.A., and Casey, L.C., April 2000 *Critical Care Clinics*). It has long been known that sepsis can be triggered by cell-wall components of Gram-negative bacteria, termed endotoxin (Takeda, K., and Akira, S. 2001 *Genes to Cells* 6:733-742).

Endotoxin is also associated with the development and progression of asthma, as well as other types of airway disease (Arbour, N.C. et al., 2000 Nature Genetics 25:187-191). In asthma and airway disease, endotoxin is believed to influence

10

15

20

25

30



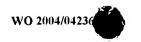
pathophysiological effects of air pollution (Arbour, N.C. et al., 2000 Nature Genetics 25:187-191). The incidence of asthma and airway diseases is increasing and, like sepsis, new treatments are needed. Effective therapeutics for these endotoxin-mediated diseases represent a serious unmet medical need.

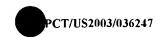
Endotoxins are composed of a lipopolysaccharide (LPS) complex, containing Lipid A and polysaccharide. The TLR4 protein has been documented (Takeda, K., and Akira, S. 2001 *Genes to Cells* 6:733-742) to recognize and bind LPS. This initiates a molecular cascade that triggers the innate immune system. Human TLR4 is known to be a homolog of the *Drosophila* protein Toll. Toll, like its human homolog, is necessary to initiate the innate immune response. Both Toll and TLR4 are known to signal through the NF-κB pathway (Medzhitov, R. et al., 1997 Nature 388:394-397).

One possible therapeutic avenue would involve inhibiting either the TLR4 gene or, more likely, the TLR4 protein, (or perhaps administration of molecules that competitively inhibit TLR4). However, this is likely to have severe and undesirable side effects. Mice strains such as C3H/HeJ and C57BL/10ScCr are unresponsive to LPS, in contrast to wild type mice, as a result of genetic defects in TLR4 (Rehli, M. et al., 2000 Journal of Biological Chemistry 275: 9773-9781). However, these strains are hypersensensitive to infection by Gram-negative bacteria (Beutler, B. 2002 Current Opinion in Hematology 9:2-10). Without a functional TLR4, and the innate immune response it triggers (which leads to an acquired immune response), these mice are unable to recognize these pathogenic invaders.

Most mammals are susceptible to septic shock. Humans, chimpanzees and bonobos are alike in extreme sensitivity to LPS, and to septic shock (Veloso, D. 1996 Immunopharmacology 33(1-3): 374-376. However, it has been well established that baboons and rhesus monkeys are resistant to septic shock, even when confronted with very high levels of LPS (Redl, H. et al., 1993 Immunobiology 187:330-345; Veloso, D. 1996 Immunopharmacology 33(1-3): 374-376). In fact, all the Old World monkeys may share this resistance to high levels of endotoxin-induced septic shock. Yet, in baboons and rhesus the innate immune response is known to be essentially the same as that of humans. Thus, baboons and rhesus have developed some mechanism for resistance to septic shock that does not interfere with innate immunity.

Because TLR4 protein is involved in septic shock, the inventors reasoned that differences in septic shock sensitivity between humans and baboons might be the





result of subtle differences in the TLR4 protein. Thus, information about the specific amino acid replacements that occurred during evolution could provide unparalleled insights into the mechanism by which baboons and rhesus monkeys resist LPS-induced septic shock while maintaining functional innate immunity.

Published *TLR4* sequences from human (GenBank AF177765, XM\_057452, U88880, and U93091), bonobo (GenBank AF179220), and baboon (GenBank AF180964) were used to design primers for polymerase chain reaction (PCR) amplification of a set of *TLR4* homologs from various primates. The primate *TLR4* homologs that were amplified and sequenced included rhesus monkey, gorilla, chimpanzee, gibbon, squirrel monkey, and capuchin. In addition, *TLR4* was amplified and sequenced from human, bonobo, and baboon and the published sequences for these species were confirmed (Seq ID: 1 to 7). As noted in Table 1, in most cases only exons 2 and 3 were sequenced (these include the full coding region of the *TLR4* gene).

15

10

5

Table 1 TLR4 Sequences

Seq. 1	Seq. 2	Seq. 3	Seq. 4	Seq. 5	Seq. 6	Seq. 7
Chimpanze	Chimpanze	Gorill	Gibbo	Rhesus	Capuchi	Squirre
e (Bonnie)	e (Dustin)	а	n	monke	n	1
Exon 3	Exons 2&3	Exons	Exons	у	Exon 3	monkey
		2&3	2&3	Exons		Exons
				2&3		2&3

20

25

These sequences were aligned and a series of molecular evolution analyses were then performed. Included in these analyses were Ka/Ks pairwise comparisons for each of these genes. Such pairwise comparisons calculate the differences between values of nonsynonymous nucleotide substitutions per nonsynonymous site (Ka) to synonymous substitutions per synonymous site (Ks). Ka values statistically significantly greater than the corresponding Ks values (Ka-Ks) strongly suggest the action of positive selection. Conversely, Ka values statistically significantly less than the corresponding Ks values (Ka-Ks) strongly suggest the action of negative selection,

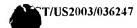
10

15

20

25

30

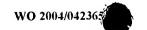


i.e., evolutionary conservation. For convenience, these pairwise comparisons are most often displayed as ratios (Ka/Ks), such that Ka/Ks >1 signifies positive selection, while Ka/Ks <1 signifies conservation.

All of these coding sequence comparisons exhibited Ka/Ks ratios less than one, some with statistical significance. This is good evidence that these are generally well-conserved, which is a commonly observed pattern. However, even well-conserved proteins can have a limited number of amino acid changes in key domains that significantly affect the function of the protein. In such cases, Ka/Ks analysis of the entire coding sequence may indicate conservation, while Ka/Ks analysis of individual domain coding regions may indicate a positively selected domain within a conserved protein. Thus, polynucleotide sequences encoding individual domains of the TLR4 protein were also subjected to Ka/Ks analysis. Two key domains are an intracellular domain responsible for signaling and an extracellular domain responsible for LPS binding.

Intracellular signaling. Ka/Ks analysis of the polynucleotide coding sequence for the TIR domain, which is the intracellular domain of TLR4 protein responsible for signaling, and which initiates the NF-κB pathway, indicates that this domain is extremely well conserved. In fact, this analysis revealed some of the lowest Ka/Ks ratios ever documented. This indicates extreme evolutionary conservation, and strongly suggests two inferences. First, this domain is a crucial one, and generally cannot tolerate amino acid replacements. Second, the signaling pathway is likely to be unchanged in all these primates. That is, regardless of differences in LPS sensitivity, the cascade initiated by the TIR domain is likely to be biochemically similar in both humans and baboons. This result thus suggests that close attention must be paid to the extracellular domain of the TLR4 protein which governs LPS recognition.

Extracellular LPS binding. LPS is thought to bind to an extracellular domain. The extracellular binding domain of TLR4 includes a number of leucine-rich repeats (LRR). These are conserved between human, bonobo, and baboon, suggesting that the basic binding mechanism is unchanged between these species. In fact, the basic LRR structure is conserved even in the Toll homolog in *Drosophila*. However, Ka/Ks analysis performed on the LPS binding domain for each primate TLR4 protein indicated the baboon LPS binding domain may be positively selected relative to the human or bonobo LPS binding domain, although there was only one nonsynonymous



15

20

25



change, thus the result was not statistically significant. This is suggestive but inconclusive evidence that the difference in septic shock sensitivity between humans and baboons results from specific amino acid replacements in the LPS binding domain.

Ka/Ks analysis of the whole protein or critical domains did not provide conclusive information about the difference in sepsis susceptibility is humans versus baboons and rhesus monkeys, so we next looked at a specific amino acid of the TLR4 gene. One human TLR4 mutation (the "human null mutation") in the extracellular ligand binding domain has been reported (Arbour, N.C. et al., 2000 Nature Genetics 25:187-191) that results in complete lack of sensitivity to LPS. The TLR4 gene from these individuals has been sequenced, and is available from GenBank (GenBank Acc. #1777766). Like baboons and rhesus, such individuals are resistant to septic shock (Arbour, N.C. et al., 2000 Nature Genetics 25:187-191). However, humans who are homozygous for this mutation have compromised immune systems and, like the C3H/HeJ and C57BL/10ScCr mice, LPS does not trigger innate immunity leaving them prone to serious Gram-negative bacterial infections.

The human null mutation is replacement of Asp299 by Gly299. Clearly, such a replacement results in substantial steric changes, leading to the loss of function observed in individuals with this mutation. Importantly, Asp299 is conserved in all mammalian TLR4s for which coding sequence data are available (except as noted below), even as phylogenetically distant from humans as mouse and rat. Such extensive conservation implies strong functional importance: this site does not generally tolerate amino acid substitutions. Importantly, however, we found that baboons and rhesus monkeys, and probably all Old World monkeys, do have an amino acid replacement at this site (Asp299 to Asn299).

		Amino acid 299	Septic Shock	Innate Immunity
30	Humans, most mammals, Drosophila	Asp	+	+
	Human null mutation	Gly	-	-
	Old World monkeys (Rhesus,			•
	baboons, etc.)	Asn	-	+

35

It is clear that inhibiting TLR4 function completely is not a viable therapeutic approach because it results in too great an impairment of the immune response.

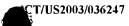
10

15

20

25

30



Similarly, modeling a therapeutic based on the human TLR4 Gly299 mutation also would result in susceptibility to Gram-negative bacterial infections. In contrast, baboons both resist septic shock and are fully capable of recognizing and addressing Gram-negative bacterial infections. This leads to a novel approach for developing therapeutic agents for treatment of endotoxin-mediated diseases of humans.

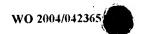
## SUMMARY OF THE INVENTION

The subject invention comprises a method of identifying an evolutionarily meaningful nucleotide change in a first primate's polynucleotide wherein the first primate's polynucleotide may be associated with a physiological condition that is present or enhanced in the first primate relative to a second primate, comprising the steps of: (a) comparing polynucleotide sequences of the first primate with corresponding polynucleotide sequences of the second primate to identify a polynucleotide that has been overall negatively selected in the first and second primates; and (b) identifying in the first primate's overall negatively selected polynucleotide, an evolutionarily meaningful nucleotide change, whereby the nucleotide change in the first primate's negatively selected polynucleotide may be associated with the physiological condition in the first primate.

The evolutionarily meaningful nucleotide change is a nonsynonymous nucleotide change in an otherwise conserved polynucleotide that is or is believed to be associated with a physiological condition. The analysis of the polynucleotides to determine whether they are negatively selected or conserved can be carried out by any method known in the art, but preferably is accomplished by a KA/KS-type analysis as described herein.

A nucleotide change in a primate's negatively selected polynucleotide can be correlated with a physiological condition in the primate by analyzing the functional effect of the presence or absence of the identified nonsynonymous nucleotide change in a model *in vivo*, *ex vivo* or *in vitro* system using methods known in the art.

In one embodiment, the first primate is an Old World monkey, the second primate is Homo sapiens, and the negatively selected polynucleotide is TLR4 polynucleotide. The evolutionarily meaningful nucleotide change in the TLR4 polynucleotide results in an Asp299 in the human and an Asn299 in the Old World monkey, e.g., baboon and rhesus monkey.

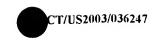


15

20

25

30



The present invention also relates to methods to develop agents for treating endotoxin-mediated diseases of humans, such that the therapeutic agent interacts with the human TLR4 polypeptide extracellular domain in such a way as to attain a response resembling that of the Old World monkeys, specifically baboons and rhesus monkeys.

It has been suggested (Arbour, N.C. et al., 2000 Nature Genetics 25:187-191) that the region of the TLR4 receptor around residue 299 has an α-helical structure. The substitution of the glycine residue (as found in the human null mutant) for the aspartic acid residue likely disrupts the 3-D structure of this helix in a catastrophic manner. However, the asparagine residue found at position 299 in the baboon and rhesus sequences is a biochemically conservative replacement compatible with the helical structure. This evolutionarily tolerated, structurally-conservative replacement allows baboons and rhesus monkeys (and most likely, all the Old World monkeys) to modulate the interaction with LPS, such that Gram-negative bacteria still trigger the innate immune response in such a way that the known resistance of both baboon and rhesus to extremely high levels of LPS is achieved. An alignment of TLR4 protein sequences for the region of the protein that flanks this residue (from a number of mammalian species) is shown in Figure 1.

It is believed that the Asp to Asn amino acid replacement at position 299 confers resistance to septic shock in baboons. Therefore, transgenic mice whose *TLR4* encodes the Asn replacement, when compared to controls (transfected with 'normal' mouse *TLR4*), should show that the experimental transgenics exhibit an increased resistance to septic shock, i.e., they will tolerate much higher levels of LPS and/or live bacteria.

The insight of the invention described herein is that the Asp299 residue is critical to initiation of the LPS-triggered cascade that leads to endotoxin-mediated diseases, such as septic shock and asthma and other airway disorders. In one preferred embodiment, a therapeutic agent is developed that, when administered, causes human TLR4 to react to exposure to endotoxin in the way baboon or rhesus monkey TLR4 does. Accordingly, in one aspect, a method is provided whereby a peptide therapeutic agent could be isolated. Such a peptide would reduce access by LPS to the key amino acid of TLR4 determining septic shock, Asp299. If delivered during an episode of acute septic shock, the peptide should "derail" the cascade that is initiated when LPS and/or live bacteria encounters the human TLR4 protein. Such a

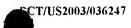
10

15

20

25

30



peptide agent can be easily tested in rodent models. Successful demonstration of protection of such models from septic shock would pave the way to similar human trials of such peptide agent. Because such a peptide would only be administered during acute episodes of septic shock, possible problems stemming from repeated administration and subsequent sensitization would be minimized. Those skilled in the art can easily determine the optimal length and amino acid composition of such therapeutic peptide, which can be further refined by testing in rodent models, using methods known in the art.

In another embodiment, a therapeutic peptide could be designed that had the same sequence as the region surrounding Asp299. Such a therapeutic could be useful as a decoy to bind to LPS and reduce the amount of LPS available to bind to the TLR4 protein.

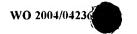
In another embodiment, an antibody or portion of an antibody could be isolated or designed that could attenuate access by endotoxin to Asp299 such that the endotoxin-mediated cascade is reduced. In a preferred embodiment, the antibody or fragment thereof is directed to an epitope that includes the Asp299 residue; the epitope preferably is an amino acid segment of 10 or less residues containing the Asp299 residue.

In another embodiment, a small molecule is identified that will reduce access to the critical Asp299 residue, or modulate the interaction of endotoxin with the Asp299 residue such that the cascade leading to endotoxin-mediated disease is modulated.

Another embodiment is to use the method disclosed herein to develop a therapeutic agent to treat human asthma. Small molecules could be designed or identified by screening libraries of small molecules that interact with Asp299 of the TLR4 polypeptide or the region containing the Asp299 polypeptide. Such therapeutic agents could be used to ameliorate the severity of asthmatic episodes. It is also likely that some therapeutic agents identified using the methods of this invention could be administered on a regular basis to reduce the effects of chronic airway diseases.

Using the teachings provided herein, persons skilled in the art will recognize that therapeutics can be developed for other diseases involving LPS – TLR4 protein interactions.

All references cited herein are each incorporated herein by reference in their entirety.

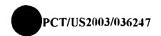


15

20

25

30



## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an alignment of TLR4 protein sequences for the region of the protein that flanks the Asp299 residue from a number of mammalian species. Amino acid residues are shown in the single letter IUPAC code. Residues that are identical in all species examined are shown in bold. Dashes have been introduced (where insertions or deletions have occurred) to maximize the alignment. The critical residue (human Asp299, baboon Asn299) is shown in lower case. Note that this Asp residue is conserved in all mammal species examined, with the exception of the biochemically-conservative Asn replacement in the Old World monkeys baboon and

Figure 2 is the nucleotide sequence for a first chimpanzee's ("Bonnie") TLR4 exon 3.

Figure 3 is the nucleotide sequence for a second chimpanzee's ("Dustin") TLR4 exons 2 and 3.

rhesus (and, importantly, the non-functional human null mutant).

Figure 4 is the nucleotide sequence for gorilla TLR4 exons 2 and 3.

Figure 5 is the nucleotide sequence for gibbon TLR4 exons 2 and 3.

Figure 6 is the nucleotide sequence for rhesus monkey TLR4 exons 2 and 3.

Figure 7 is the nucleotide sequence for capuchin TLR4 exon 3.

Figure 8 is the nucleotide sequence for squirrel monkey TLR4 exon 3.

## DETAILED DESCRIPTION OF THE INVENTION

The subject invention relates to a method of identifying a nucleotide change in a TLR4 polynucleotide sequence of an Old World monkey wherein such change may be associated with reduced sensitivity to Gram-negative bacterial infection. This method involves the comparison of the TLR4 polynucleotide sequence from the Old World monkey with corresponding TLR4 polynucleotide sequence of a human to identify a polynucleotide change in said Old World monkey's TLR4 sequence that is evolutionarily meaningful. The evolutionarily meaningful change may then be associated with reduced sensitivity to Gram-negative bacterial infection. In particular, the evolutionarily meaningful change is from Asp299 in the human to Asn299 in the rhesus monkey or baboon.

The subject invention also includes a method of identifying a therapeutic agent that reduces sensitivity to Gram-negative bacterial infection. This method comprises

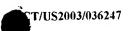
10

15

20

25

30



(a) contacting candidate agents with human TLR4 polypeptide; and (b) identifying a therapeutic agent that interacts with the TLR4 polypeptide to substantially reduce sensitivity to Gram-negative bacterial infection. The agent interaction with TLR4 polypeptide preferably occurs at Asp299.

The therapeutic agent identified according to the subject invention can be used in the treatment of sepsis, severe sepsis, septic shock, asthma or other respiratory ailments, in humans or non-human primates.

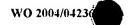
The subject invention also concerns a number of novel *TLR4* polynucleotide sequences as provided in Figures 2-9, and their deduced TLR4 polypeptide sequences. These sequences are chimpanzee, gorilla, gibbon, rhesus monkey, capuchin, squirrel monkey and baboon. They are useful in the identification of evolutionarily meaningful nucleotide changes in other primate *TLR4* polynucleotides. Their polynucleotide or polypeptide sequences may also be useful in the design of candidate therapeutic agents according to the subject invention.

The practice of the present invention employs, unless otherwise indicated, conventional techniques of molecular biology, genetics, and assay development, which are within the skill of the art. Such techniques are explained fully in the literature, such as: "Molecular Cloning: A Laboratory Manual", second edition (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Current Protocols in Molecular Biology" (F.M. Ausubel et al., eds., 1987); "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994)

#### **Definitions**

As used herein, a "polynucleotide" refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides, or analogs thereof. This term refers to the primary structure of the molecule, and thus includes double- and single-stranded DNA, as well as double- and single-stranded RNA. It also includes modified polynucleotides such as methylated and/or capped polynucleotides. The terms "polynucleotide" and "nucleotide sequence" are used interchangeably.

As used herein, a "gene" refers to a polynucleotide or portion of a polynucleotide comprising a sequence that encodes a protein. It is well understood in the art that a gene also comprises non-coding sequences, such as 5' and 3' flanking sequences (such as promoters, enhancers, repressors, and other regulatory sequences) as well as introns.



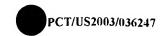
10

15

20

25

30



The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. These terms also include proteins that are post-translationally modified through reactions that include glycosylation, acetylation and phosphorylation.

The term " $K_A/K_S$ -type methods" means methods that evaluate differences, frequently (but not always) shown as a ratio, between the number of nonsynonymous substitutions and synonymous substitutions in homologous genes (including the more rigorous methods that determine non-synonymous and synonymous sites). These methods are designated using several systems of nomenclature, including but not limited to  $K_A/K_S$ ,  $d_N/d_S$ ,  $D_N/D_S$ .

The term "evolutionarily meaningful change" refers to one or more nonsynonymous nucleotide change(s) or corresponding amino acid change(s) between two species that occurs in an otherwise conserved polynucleotide or polypeptide, that may be attributed to a positive selective pressure, and which is or is believed to be associated with a physiological condition. A conserved polynucleotide can be identified by methods known in the art including a KA/KS-type analytical method. Typically, a K<sub>A</sub>/K<sub>S</sub> ratio less than about 1.0, more preferably less than about 0.75, and most preferably less than about 0.5 indicates the action of negative selection. presence of a nonsynonymous nucleotide change in such a conserved polynucleotide (i.e., containing no other nucleotide changes or only synonymous nucleotide changes) is considered to be an evolutionarily meaningful change. The phrase "associated with a physiological condition" means that the nonsynonymous nucleotide change has been observed in individuals to result in the physiological condition at issue, has been shown to be involved in a molecular mechanism related to the physiological condition, and/or occurs in a location of the gene that is relevant to a protein function that is essential to the occurrence of the physiological condition. For example, as discussed herein, a nonsynonymous nucleotide change in the baboon or rhesus monkey gene is believed to be associated with the physiological condition of enhanced resistance to the endotoxin-mediated response.

The term "positive evolutionarily meaningful change" means an evolutionarily meaningful change in a particular species that results in an adaptive change that is positive as compared to other related species. An example of a positive evolutionarily meaningful change is a change that has resulted in reduced sensitivity to the LPS mediated response.

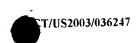
10

15

20

25

30



The term "resistant" means that an organism exhibits an ability to avoid, or diminish the extent of, a disease condition and/or development of the disease, preferably when compared to non-resistant organisms.

The term "susceptibility" means that an organism fails to avoid, or diminish the extent of, a disease condition and/or development of the disease condition, preferably when compared to an organism that is known to be resistant.

It is understood that resistance and susceptibility vary from individual to individual, and that, for purposes of this invention, these terms also apply to a group of individuals within a species, and comparisons of resistance and susceptibility generally refer to overall, average differences between species, although intra-specific comparisons

may be used.

The term "nucleotide change" refers to nucleotide substitution, deletion, and/or insertion, as is well understood in the art.

The term "agent", as used herein, means a biological or chemical compound such as a simple or complex organic or inorganic molecule, a peptide, a protein or an oligonucleotide that modulates the function of a polypeptide. A vast array of compounds can be synthesized, for example oligomers, such as oligopeptides and oligonucleotides, and synthetic organic and inorganic compounds based on various core structures, and these are also included in the term "agent". In addition, various natural sources can provide compounds for screening, such as plant or animal extracts, and the like. The term "agent" can include or exclude antibodies or fragments thereof. Compounds can be tested singly or in combination with one another.

The term "therapeutic agent" as used herein means an agent as described above used to treat a disease or condition.

The term "to modulate function" of a polypeptide means that the function of the polypeptide is altered in the presence of an agent compared to the absence of the agent. Modulation may occur on any level that affects function. Modulation of a polypeptide function may be direct or indirect, and measured directly or indirectly.

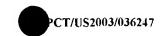
A "function of a polypeptide" includes, but is not limited to, conformation, folding (or other physical characteristics), binding to other moieties (such as ligands), activity (or other functional characteristics), and/or other aspects of protein structure or functions. For example, an agent that acts on a polypeptide and affects its conformation, folding (or other physical characteristics), binding to other moieties

15

20

25

30



(such as ligands), activity (or other functional characteristics), and/or other aspects of protein structure or functions is considered to have modulated polypeptide function. The ways that an effective agent can act to modulate the function of a polypeptide include, but are not limited to 1) changing the conformation, folding or other physical characteristics; 2) changing the binding strength to its natural ligand or changing the specificity of binding to ligands; and 3) altering the activity of the polypeptide.

The term "to modulate the endotoxin or LPS mediated response" means that the function of the TLR 4 polypeptide is altered in the presence of an agent compared to the absence of the agent. The modulation reduces the clinical symptoms of sepsis, severe sepsis, or septic shock, including central nervous system dysfunction, adult respiratory distress syndrome, liver failure, acute renal failure, disseminated intravascular coagulation, and the like. Preferably, these symptoms are reduced in increasing preference by at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% and 100%, as measured by standard clinical indicators or assays for those symptoms as is known in the art. Modulation can be detected in other ways as well, including reduced affinity or altered kinetics of LPS binding to the TLR4 extracellular domain, or reduced signaling to a downstream effector of TLR4 in the LPS mediated response. Preferably, the agent modulation results in the attenuation of the LPS mediated response to the degree that the human TLR4 interacts with the LPS in the same manner as the Old World monkey TLR4.

Further, "modulation of endotoxin or LPS mediated response" means significant reduction or attenuation of the response whereby the clinical symptoms of sepsis, severe sepsis or septic shock are reduced as indicated above. It does not refer to abrogation or elimination of the response; i.e., innate immunity remains intact. The therapeutic agent of the subject invention is one which interacts directly or indirectly with the Asp299 residue such that the activation of the NF-kB pathway and the clinical symptoms of sepsis, severe sepsis and septic shock are attenuated.

The term "target site" means a location in a polypeptide which can be a single amino acid and/or is a part of, a structural and/or functional motif, e.g., a binding site, a dimerization domain, or a catalytic active site. Target sites may be useful for direct or indirect interaction with an agent, such as a therapeutic agent.

The term "positively selected" means an evolutionarily significant change in a particular organism, species, subspecies, variety, cultivar or strain that results in an adaptive change that is positive as compared to other related organisms. An example

10

20

25

30



of a positive evolutionarily significant change is a change that has resulted in enhanced yield in crop plants. As stated above, positive selection is indicated by a  $K_A/K_S$  ratio greater than 1.0. With increasing preference, the  $K_A/K_S$  value is greater than 1.25, 1.5 and 2.0.

"TLR4" as used herein refers to the polynucleotide encoding the Toll like receptor 4 polypeptide.

"TLR4" as used herein refers to the polypeptide encoded by TLR4 polynucleotide.

"LPS" as used herein refers to lipopolysaccharide and is used interchangeably with the word "endotoxin".

## General Methods of the Invention

The general method of the invention is as follows. Briefly, the *TLR4* polynucleotide sequences are obtained from a human source and a number of nonhuman primate sources. They are compared to one another to determine whether the TLR4 polynucleotide is conserved or negatively selected. Then, having determined that the polynucleotide is conserved, the *TLR4* polynucleotide sequences from the human and other primates are analyzed to identify any nonsynonymous or evolutionarily meaningful nucleic acid differences. The *TLR4* sequences from each species are then characterized in terms of whether they do or do not correlate with decreased sensitivity to Gram-negative bacterial infection for that species, thereby indicating those evolutionarily meaningful changes that could be or are associated with the decreased sensitivity to LPS. This method resulted in the identification of Asn299 as the critical amino acid found in baboons and rhesus monkeys that confers attenuated sensitivity to LPS in those species as compared to humans which have Asp299.

U.S. Serial No. 10/100,422, filed March 18, 2002, and incorporated herein in its entirety by reference, describes in detail a number of methods useful in sequencing homologous polynucleotide sequences from primates, and methods for identification of evolutionarily meaningful changes in polynucleotides that can be correlated with a particular physiological condition in humans or in non-human primates.

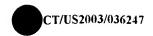
Determination of evolutionarily meaningful changes first requires a determination as to whether the polynucleotide at issue is negatively selected. Any of several different molecular evolution analyses or  $K_A/K_S$ -type methods can be

15

20

25

30



employed to determine whether the human gene sequences and the non-human primate polynucleotide are conserved or negatively selected. Kreitman and Akashi (1995) Annu. Rev. Ecol. Syst. 26:403-422; Li, Molecular Evolution, Sinauer Associates, Sunderland, MA, 1997. For example, negative selection on proteins (i.e., molecular-level conservation) can be detected in protein-coding genes by pairwise comparisons of the ratios of nonsynonymous nucleotide substitutions per nonsynonymous site (KA) to synonymous substitutions per synonymous site (KS) (Li et al., 1985; Li, 1993). Any comparison of KA and KS may be used, although it is particularly convenient and most effective to compare these two variables as a ratio. Negatively selected polynucleotides are identified by having a KA/KS ratio of less than 1.0, preferably less than 0.75 and more preferably less than 0.5. Preferably, the  $K_A/K_S$ analysis by Li et al. is used to carry out the present invention, although other analysis programs that can detect negatively selected genes between species can also be used. Li et al. (1985) Mol. Biol. Evol. 2:150-174; Li (1993); see also J. Mol. Evol. 36:96-99; Messier and Stewart (1997) Nature 385:151-154; Nei (1987) Molecular Evolutionary Genetics (New York, Columbia University Press). The KA/KS method, which comprises a comparison of the rate of non-synonymous substitutions per nonsynonymous site with the rate of synonymous substitutions per synonymous site between homologous protein-coding region of genes in terms of a ratio, can be used to identify conserved polynucleotides. A synonymous ("silent") substitution is one that, owing to the degeneracy of the genetic code, makes no change to the amino acid sequence encoded; a nonsynonymous substitution results in an amino acid replacement. The extent of each type of change can be estimated as  $K_A$  and  $K_S$ , respectively, the numbers of synonymous substitutions per synonymous site and nonsynonymous substitutions per non-synonymous site. Calculations of K<sub>A</sub>/K<sub>S</sub> may be performed manually or by using software. An example of a suitable program is MEGA (Molecular Genetics Institute, Pennsylvania State University).

For the purpose of estimating K<sub>A</sub> and K<sub>S</sub>, either complete or partial human and primate protein-coding sequences are used to calculate total numbers of synonymous and non-synonymous substitutions, as well as non-synonymous and synonymous sites. The length of the polynucleotide sequence analyzed can be any appropriate length, but is preferably at least 60 nucleotides in length. Preferably, the entire coding sequence is compared, in order to determine overall conservation. Publicly available

10

15

20

25

30



computer programs, such as Li93 (Li (1993) J. Mol. Evol. 36:96-99) or INA, can be used to calculate the  $K_A$  and  $K_S$  values for all pairwise comparisons.

As indicated above, conservation is indicated by the  $K_A/K_S$  ratio being less than about 1.0, more preferably less than about 0.75, and most preferably less than about 0.5. Preferably, statistical analysis is performed on all decreased  $K_A/K_S$  ratios, including, but not limited to, standard methods such as Student's *t*-test and likelihood ratio tests described by Yang (1998) *Mol. Biol Evol.* 37:441-456.

All methods for calculating K<sub>A</sub>/K<sub>S</sub> ratios are based on a pairwise comparison of the number of nonsynonymous substitutions per nonsynonymous site to the number of synonymous substitutions per synonymous site for the protein-coding regions of homologous genes from related species. Each method implements different corrections for estimating "multiple hits" (i.e., more than one nucleotide substitution at the same site). Each method also uses different models for how DNA sequences change over evolutionary time. Thus, preferably, a combination of results from different algorithms is used to increase the level of sensitivity for detection of negatively-selected genes and confidence in the result.

As discussed above, the foregoing methods resulted in the identification of an evolutionarily meaningful nucleotide change in the conserved *TLR4* polynucleotide. The Asn299 of baboons and rhesus monkeys was found to attenuate LPS sensitivity in those species relative to humans which have Asp299. This information can be useful in the development of agents that interact with human TLR4 Asp299 in such a manner so as to attenuate activation of the NF-kB pathway by LPS, therefore aiding in the treatment of sepsis, severe sepsis and septic shock.

An agent is designed or identified that will interact with human TLR4 protein in such a way that the agent modulates the human endotoxin mediated response. Preferably, the agent causes the human TLR4 to interact with LPS in a manner that is similar to that of baboon or rhesus TLR4 interaction with LPS. Such agents can be peptide, protein, organic molecules, or aptamers, or whatever agent can have the specific effect.

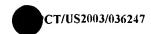
Generally, techniques of combinatorial chemistry can be used to generate numerous permutations of agent candidates to be screened for effectiveness in reducing access of LPS to Asp299 of TLR4. Those of skill in the art can devise and/or obtain suitable agents for testing. In general, screening can be performed by adding an agent to a sample of isolated TLR4 or appropriate cells expressing TLR4

15

20

25

30



and monitoring the effect, that is, modulation of the cascade known to lead to endotoxin-mediated disease. The experiments preferably include a control sample which does not receive the candidate agent. Differences between treated and untreated cells indicate effects attributable to the candidate agent. Optimally, a greater effect is seen in the presence of the candidate agent than in the absence of the candidate agent. Hoshino, K., Takeuchi, O., Kawai, T., et al. (1999). J. Immunol. 162, 3749-3752. describe a typical assay for measuring the NF-kB pathway that is known to lead to endotoxin-mediated disease.

The screening methods for agents that interact with TLR4 polypeptide can be carried out *in vitro*, *ex vivo* or *in vivo* using TLR4 protein or polypeptide or extracellular fragment thereof, or NF-kB pathway models known in the art.

In an example for an assay for an agent that binds to TLR4 polypeptide, an affinity column is prepared with purified human TLR4 or a synthetically prepared peptide of a small region of TLR4 containing Asp299. The affinity column is then used to screen a library of compounds (libraries of compounds include, but are not limited to, peptides, aptamers, small molecules, etc.) which have been appropriately labeled. Suitable labels include, but are not limited to fluorochromes, radioisotopes, enzymes and chemiluminescent compounds. The unbound and bound compounds can be separated by washes using various conditions known to those skilled in the art. In addition to affinity columns, there are other techniques, such as measuring the fluorescence anisotropy of a protein which will change upon binding another molecule. For example, a BIAcore assay using a sensor chip (supplied by Pharmacia Biosensor, Stitt et al. (1995) Cell 80:661-670.) that is coupled to TLR4 or a peptide of TLR4 containing Asp299 may be performed to determine the binding activity of different agents.

It is also understood that the in vitro screening methods of this invention include structural, or rational, drug design, in which the amino acid sequence surrounding Asp299, three-dimensional atomic structure or other property (or properties) of the amino acid sequence surrounding Asp299 provides a basis for designing an agent which is expected to bind to TLR4 in such a way as to reduce access by endotoxin to Asp299.

The screening methods described above represent primary screens, designed to detect any agent that may bind to TLR4 and/or exhibit activity that modulates the function of TLR4. The skilled artisan will recognize that secondary tests will likely

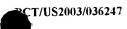
10

15

20

25

30



be necessary in order to evaluate an agent further. For example, a secondary screen may comprise testing the agent(s) in a mouse model or other animal models for effect in reducing the endotoxin-mediated cascade leading to disease.

The invention also includes agents identified by the screening methods described herein.

## Peptide agents

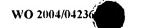
A peptide agent can be isolated by screening a library of randomly synthesized peptides for peptides that bind to residue 299, using methods known in the art, for example, as described in Dower, William J; Cwirla, Steven E; Barrett, Ronald W, "Peptide library and screening systems." *Biotechnol Advances* 1996 14(4):490. Peptides found to bind to TLR4 in such a way that the endotoxin does not interact directly with Asp 299 are selected. In one embodiment, the peptide library contains randomly synthesized peptides of at least 20 amino acids in length, preferably less than 50 amino acids in length. In another embodiment, the peptide library contains randomly synthesized peptides of between 15 and 20 amino acids in length. In a third embodiment, the peptide library contains peptides between 10 and 14 amino acids in length. In a fourth embodiment, the peptide library contains peptides between 1 and 9 amino acids in length. Peptides found to bind TLR4 at or near residue 299 are then subjected to secondary screens in vitro for effect on the binding of LPS to TLR4 and prevention or attenuation of TLR4 signaling.

#### **Aptamers**

Similar to that described above, the subject invention includes methods to identify an aptamer agent by screening a library of randomly synthesized single-stranded nucleotides, using methods known in the art, for example, as described in Bell, C.; Lynam, E.; Landfair, D.J.; Janjic, N.; and Wiles, M.E., "Oligonucleotides NX1838 inhibits VEGF165-mediated cellular responses in vitro", *In Vitro Cell Dev. Biol. Anim.* 1999 Oct:35(9):533-42.

## Small molecule agents

One method for developing such a small molecule agent would be to use 3dimensional modeling of the secondary and tertiary structure of the region of TLR4 that surrounds the critical Asp299 residue. Potential small molecules can be 'docked'





in silico, in order to identify a close fit that will inhibit access to the aspartic acid residue or have the desired structural affect. Likely candidates can then be tested in vivo in rodent models. Small molecule therapeutics could also be identified from libraries of small molecules through the use of assays that screen for compounds that bind to the region of TLR4 containing Asp299. In a similar fashion, it may be possible to design or screen for a small molecule that does not completely block access to the Asp299 residue, but rather modulates the kinetics of LPS binding in this region in such a way that it more closely resembles the kinetics of LPS binding to baboon TLR4 protein.

10

15

20

25

30

Therapeutic compositions that comprise agents

As described herein, agents can be screened for their capacity to modulate the LPS mediated NF-kB pathway.

Various delivery systems are known in the art that can be used to administer agents identified according to the subject methods. Such delivery systems include aqueous solutions, encapsulation in liposomes, microparticles or microcapsules or conjugation to a moiety that facilitates intracellular admission.

Therapeutic compositions comprising agents may be administered parenterally by injection, although other effective administration forms, such as intra-articular injection, inhalant mists, orally-active formulations, transdermal iontophoresis or suppositories are also envisioned. The carrier may contain other pharmacologically-acceptable excipients for modifying or maintaining the pH, osmolarity, viscosity, clarify, color, sterility, stability, rate of dissolution, or odor of the formulation. The carrier may also contain other pharmacologically-acceptable excipients for modifying or maintaining the stability, rate of dissolution, release or absorption of the agent. Such excipients are those substances usually and customarily employed to formulate dosages for parenteral administration in either unit dose or multi-dose form.

Once the therapeutic composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or dehydrated or lyophilized powder. Such formulations may be stored either in a ready to use form or requiring reconstitution immediately prior to administration. The manner of administering formulations containing agents for systemic delivery may be via subcutaneous, intramuscular, intravenous, intranasal or vaginal or rectal suppository.

10

15

20

25

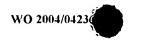
30

The amount of agent which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, which can be determined by standard clinical techniques. In addition, in vitro or in vivo assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness or advancement of the disease or condition, and should be decided according to the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems. For example, an effective amount of an agent identified according to the subject methods is readily determined by administering graded doses of the agent and observing the desired effect.

The following examples are provided to further assist those of ordinary skill in the art. Such examples are intended to be illustrative and therefore should not be regarded as limiting the invention. A number of exemplary modifications and variations are described in this application and others will become apparent to those of skill in this art. Such variations are considered to fall within the scope of the invention as described and claimed herein.

# Example 1. PCR amplification and DNA sequencing of primate TLR4 sequences.

Published *TLR4* sequences from human (GenBank AF177765, XM\_057452, U88880, and U93091), bonobo (GenBank AF179220), and baboon (GenBank AF180964) were used to design primers (by methods well-known to those skilled in the art) for polymerase chain reaction (PCR) amplification of a set of *TLR4* homologs from various primates. The primate *TLR4* homologs that were PCR amplified and DNA sequenced (by methods well-known to those skilled in the art) included rhesus monkey, gorilla, chimpanzee, gibbon, squirrel monkey, and capuchin. In addition, *TLR4* was amplified and sequenced from human, bonobo, and baboon and the published sequences for these species were confirmed (Seq ID: 1 to 7). Because exons 2 and 3 contain the full coding region of the *TLR4* gene, in most cases only exons 2 and 3 were sequenced. These sequences were aligned by methods well-known to those skilled in the art.

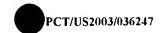


15

20

25

30



## Example 2. Ka/Ks analysis.

Ka/Ks pairwise comparisons were completed for each of these genes. Such pairwise comparisons calculate the differences between values of nonsynonymous nucleotide substitutions per nonsynonymous site (Ka) to synonymous substitutions per synonymous site (Ks). Ka values statistically significantly greater than the corresponding Ks values (Ka-Ks) strongly suggest the action of positive selection. Conversely, Ka values statistically significantly less than the corresponding Ks values (Ka-Ks) strongly suggest the action of negative selection, i.e., evolutionary conservation. For convenience, these pairwise comparisons are most often displayed as ratios (Ka/Ks), such that Ka/Ks >1 signifies positive selection, while Ka/Ks <1 signifies conservation.

All of these whole protein comparisons exhibited Ka/Ks ratios less than one. some with statistical significance. This is good evidence that these are generally wellconserved proteins, which is a commonly observed pattern. However, even wellconserved proteins can have amino acid changes in key domains that significantly affect the function of the protein. In such cases, Ka/Ks analysis of the entire coding sequence may indicate conservation, while Ka/Ks analysis of individual domain coding regions may indicate a positively selected domain within a conserved protein. Thus, polynucleotides encoding individual domains of the TLR4 protein were also subjected to analysis. Two key domains are an intracellular domain responsible for signaling and an extracellular domain responsible for LPS binding. Ka/Ks analysis was performed for the TIR domain, which is the intracellular domain of TLR4 protein responsible for signaling, and which initiates the NF-κB pathway. This analysis indicated that this domain is extremely well conserved. In fact, this analysis revealed some of the lowest Ka/Ks ratios ever documented. Ka/Ks analysis was then performed for the extracellular signaling domain of TLR. Here the result was inconclusive, in that although evidence was seen for possible positive selection on the extracellular LPS-binding domain of baboon TLR4 (relative to human TLR4), no statistical support exists for this. As a result, further analysis was performed (see Example 3).

#### Example 3. Further molecular evolutionary analysis.

Further analysis included a search for individual amino acid replacements that are either shared by, or are unique to, the human and baboon TLR4 sequences. One

15

20

25

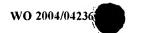
human TLR4 mutation in the extracellular ligand binding domain has been reported (Arbour, N.C. et al., 2000 Nature Genetics 25:187-191) that results in complete lack of sensitivity to LPS. Like baboons and rhesus, such individuals are resistant to septic shock. However, humans who are homozygous for this mutation have compromised immune systems and LPS does not trigger innate immunity leaving them prone to serious Gram-negative bacterial infections. The human null mutation is replacement of Asp299 by Gly299. Importantly, Asp299 is conserved in all mammalian TLR4s for which sequence is available, even as phylogenetically distant as mouse and rat, with the exception of baboon and rhesus, which have a biochemically conservative replacement amino acid replacement at this site (Asp299 to Asn299). The substitution of the glycine residue (as found in the human null mutant) for the aspartic acid residue likely disrupts the 3-D structure of this helix in a catastrophic manner. However, because the asparagine residue found at position 299 in the baboon and rhesus sequences is a biochemically conservative replacement, it is likely compatible with the helical structure. This evolutionarily tolerated, structurally-conservative replacement thus probably allows baboons and rhesus monkeys (and most perhaps, all the Old World monkeys) to modulate the interaction with LPS, such that Gramnegative bacteria still trigger the innate immune response in such a way that the known resistance of both baboon and rhesus to extremely high levels of LPS is achieved.

## Example 4. Design of a peptide therapeutic agent

A library of random peptides 20 amino acids long is synthesized and screened for peptide agents that bind to TLR4. A secondary screen then assesses peptides found in this primary screen for ability to reduce access by LPS to Asp299 and measurably reducing the LPS-mediated cascade leading to septic shock. The optimal length and amino acid composition of the therapeutic peptide agent can be refined by testing in rodent models, as would be known to one skilled in the art.

## 30 Example 5. Design and synthesis of a decoy peptide agent.

A peptide is designed that has the same sequence as the region of human TLR4 from amino acid 289 through 309. This peptide is synthesized synthetically and formulated for delivery as a therapeutic. Such a peptide therapeutic would be useful



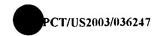
10

15

20

25

30



as a decoy to bind to LPS and reduce the amount of LPS available to bind to the TLR4 protein.

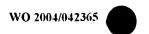
## Example 6. Design and synthesis of a small molecule therapeutic agent.

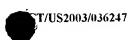
The secondary and tertiary structure of the region of TLR4 protein from amino acid 289 through 309 is modeled. Small molecule core structures are screened *in silico* for the ability to 'dock' in this region in order to identify a close fit that will inhibit access by endotoxin to the Asp299 residue either by directly or indirectly reducing access to Asp299. Likely candidates can then be tested *in vivo* in rodent models.

## Example 7. Screening for a small molecule therapeutic.

Libraries of small molecules are purchased from one of several vendors. An affinity column is prepared with purified human TLR4 or a synthetically prepared peptide of a small region of TLR4 containing Asp299. The affinity column is then used to screen a library of compounds which have been appropriately labeled. Suitable labels include, but are not limited to fluorochromes, radioisotopes, enzymes and chemiluminescent compounds. The unbound and bound compounds can be separated by washes using various conditions known to those skilled in the art. In addition to affinity columns, there are other techniques, such as measuring the fluorescence anisotropy of a protein which will change upon binding another molecule. For example, a BIAcore assay using a sensor chip (supplied by Pharmacia Biosensor, Stitt et al. (1995) Cell 80:661-670.) that is coupled to TLR4 or a peptide of TLR4 containing Asp299 may be performed to determine the binding activity of different agents.

A secondary screen is then employed to identify small molecules that reduce access by endotoxin to the Asp299 residue, or modulate the kinetics of endotoxin binding in the region containing Asp299 in such a way that it more closely resembles the kinetics of endotoxin binding to baboon TLR4 protein. For example, mammals that are susceptible to LPS-mediated response (i.e., those with Asp299) could be administered an appropriate dose of the candidate agent to determine if it attenuates the sepsis symptoms normally associated with exposure to LPS.





## Example 8. Screening antibody candidates

Antibodies (or modified antibodies or antibody fragments) are isolated/designed that bind to the extracellular region of human TLR4 such that access by bacterial LPS is diminished to Asp299. Such antibodies are directed to an epitope comprising Asp299. Preferably the epitope is 10 or less residues in length. Creation or isolation of such antibodies is understood by those skilled in the art. Uehori et al. (2003) Infect. Immun. 71(8):4238, describe antibodies to TLR4 which inhibited bacterial cell wall skeleton mediated NF-kB activation by 80%. Also see akashi et al. (2000) J. Immunol. 164:3471-75.

#### Claims:

5

10

- 1. A method of identifying a nucleotide change in a TLR4 polynucleotide sequence of an Old World monkey wherein said change may be associated with reduced sensitivity to Gram-negative bacterial infection, comprising the step of: comparing the TLR4 polynucleotide sequence of the Old World monkey with corresponding TLR4 polynucleotide sequence of a human to identify a polynucleotide change in said Old World monkey's TLR4 sequence that is evolutionarily meaningful, whereby said evolutionarily meaningful change may be associated with reduced sensitivity to Gram-negative bacterial infection.
  - 2. The method of claim 1 wherein the Old World monkey is selected from the group consisting of rhesus monkey and baboon.
- 15 3. The method of claim 2 wherein the evolutionarily meaningful change is from Asp299 in the human to Asn299 in the rhesus monkey or baboon.
  - 4. The method of claim 3, wherein the evolutionarily meaningful change is associated with the reduced sensitivity to Gram-negative bacterial infection by the step comprising:

analyzing the functional effect of the Asp299Asn change in a model system.

5. The method of claim 4, wherein said model system is in vivo, ex vivo or in vitro.

25

20

- 6. A method of identifying a therapeutic agent that reduces sensitivity to Gramnegative bacterial infection, comprising:
  - (a) contacting candidate agents with human TLR4 polypeptide; and
- (b) identifying a therapeutic agent that interacts with the TLR4
   30 polypeptide to substantially reduce sensitivity to Gram-negative bacterial infection.
  - 7. The method of claim 6 wherein said interaction with TLR4 polypeptide occurs at Asp299.

10

15

25



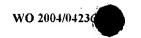
- 8. The method of claim 6, wherein said substantial reduction in sensitivity to Gram-negative bacterial infection is determined by an indicator selected from the group consisting of:
- (a) elimination or substantial reduction in host systemic inflammatory response to LPS in a human, non-human primate, or suitable animal model; and
- (b) elimination or reduced severity of central nervous system dysfunction, adult respiratory distress syndrome, liver failure, acute renal failure, and/or disseminated intravascular coagulation in a human, non-human primate, or suitable animal model.

9. A method for treating sepsis, severe sepsis or septic shock in a primate, comprising:

administering to a primate in need thereof an effective dose of a therapeutic agent identified according to the method of claim 6.

10. A method for treating asthma in a primate, comprising: administering to a primate in need thereof an effective dose of a therapeutic agent identified according to the method of claim 6.

- 20 11. A therapeutic agent identified according to the method of claim 6.
  - 12. A composition comprising a polynucleotide selected from the group consisting of chimpanzee *TLR4* polynucleotide (SEQ. ID. NO. \_\_), gorilla *TLR4* polynucleotide (SEQ. ID. NO. \_\_), rhesus monkey *TLR4* polynucleotide (SEQ. ID. NO. \_\_), capuchin *TLR4* polynucleotide (SEQ. ID. NO. \_\_), squirrel monkey *TLR4* polynucleotide (SEQ. ID. NO. \_\_), and baboon *TLR4* polynucleotide (SEQ. ID NO. \_\_), and baboon *TLR4* polynucleotide (SEQ. ID NO. ).
- 13. A composition comprising a polypeptide selected from the group consisting of chimpanzee TLR4 polypeptide (SEQ. ID. NO. \_\_), gorilla TLR4 polypeptide (SEQ. ID. NO. \_\_), gibbon TLR4 polypeptide (SEQ. ID. NO. \_\_), rhesus monkey TLR4 polypeptide (SEQ. ID. NO. \_\_), capuchin TLR4 polypeptide (SEQ. ID. NO. \_\_), squirrel monkey TLR4 polypeptide (SEQ. ID NO. \_\_), and baboon TLR 4 polypeptide (SEQ. ID NO. ).





Human	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Human null	CNLTIEEFRLTYLD-YYLDgIIDLFNCLANASSFSL
Chimpanzee	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Bonobo	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Gorilla	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Orangutan	CNLTIEEFRLAYLD-YYLDdIIDLFNCLANVSSFSL
Gibbon	CNLTIEEFRLTYLD-YYLDdIIDLFNCLANASSFSL
Baboon	CNLTIEEFRLTYLD-YYLDnIIDLFNCLANASSFSL
Rhesus	CNLTIEEFRLTYLD-YYLDnIIDLFNCLANASSFSL
Horse	HNLTIEEFRLAYIDNYSSKdSIDLLNCLADISKISL
Cow	CNLTIEQFRIAYLDKFSGDd-TDLFNCLANVSVISL
Cat	CNLIIEKFRIAYFDKFS-EdAIDSFNCLANVSTISL
Dog	CNLTIEKFRIAYFDSFS-KdTTNLFNQLVNISAISL
Hamster	CKVTIEEFRFTYANEFS-EdITD-FDCLANVSAMSL
Rat	CNVSIDEFRLTYINHFS-Ddiyn-LnCLANISAMSF
Mouse	CDVTIDEFRLTHTNDFS-DdI-VKFHCLANVSAMSL

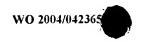
Figure 1.

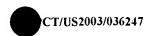
2/9

T/US2003/036247

Baboon CDS

GTGGTTCCTAACATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACA TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA TAGCTTCCTCCGTTTTCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG CTGTGGAGACAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA AGAACTTAATGTGGCTCACAATCTTATCCAGTCTTTCAAATTACCTGAGTATTTTTCTAATC CTTGCAGGTTCTACATCAAATGCCCCTACCCAATCTCTCTTTAGACCTGTCCCTGAACCCTA TAAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAG CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAACTTGGAAGAGTTTGACAAATCT GCTCTGGAGGGATTGTGCAATTTGACCATTGAAGAATTCCGATTAACATACTTAGACTACT ACCTCGATAATATTATTGACTTATTTAATTGTTTGGCAAATGCTTCTTCATTTTCCCTGGTG AGTGTGAATATTAAAAGGGTAGAAGACTTTTCTTATAATTTCAGATGGCAACATTTAGAAT TAGTTAACTGTAAATTTGAACAGTTTCCCACATTGGAACTCGAATCTCTCAAAAGGCTTAC TTTCACTGCCAACAAAGGTGGGAATGCCTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTATTACCATGGGTTCAAACTT CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAGATGAGT CAATTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC CACAGTTGCTTTCAATGGCATTTTCGATGGCTTGCTCAGAGTCTTAAAAATGGCT GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGATCTGAAAAACTTGACCT TCCTGGACCTCTCAGTGTCAACTGGAGCAGTTGTCTCCAACAGCATTTGACACACTCAA CAAGCTTCAGGTACTAAATATGAGCCACAACAACITCTTTTCATTGGATGTGTTTCCTTAT AAGTGTCTGCCCTCCAGGTTCTCGATTACAGTCTCAATCACATAATGACTTCCAAAA ACCAGGAACCTCAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT TGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCAGCTCTTG GTGGAAGCTGAACGAATGGAATGTGCAACACCTTCAGATAAACAGGGCATGCCTGTGCTG AGTGTGAATATTACCTGTCAGATGAATAAGACCATCATTGGTGTGTCTGTGTTCAGTGTGC TTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT GGCTGCATAAAGTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGCC AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCC TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCAAACATCA TCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCA GAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCGT GCAGGCATAATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGTG GAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTAGGG CAGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGTTGGATGGCAGATCGTGGAATCCA **GAAGAACAGTAG** 



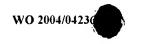


Bonobo

GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC TCCCCTTCTCAACCAAGAACCTGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC  ${\tt CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG}.$ CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA AGAACTTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC CTTGCGGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA TGAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAA CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAGAAAACTTGGAAAAGTTTGACAAATCT GCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACTACT ACCTCGATGATATTATTGACTTATTTAATTGTTTGACAAATGTTTCTTCATTTTCCCTGGTG AGTGTGACTATTAAAAGCGTAAAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAAT TAGTTAAGTGTAAATTTGGACAGTTTCCCACATTGAAACTCAAATCTCTCAAAAGGCTTAC TTTCACTTCCAACAAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGGTGTTATTACCATGAGTTCAAACTT CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC CAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAGTCTCGAAGTCTTGAAAATGGCT GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGACCT CAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTAT **AAGTGTCTGAACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAA** AACAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT TGCTTGTACTTGTGAACACCAAAGTTTCCTGCAATGGATCAAGGACCAGAGGCAGCTCTTG GTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCTG AGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCGGTCCTCAGTGTGC TTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT GGCTGCATAAAGTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGCC AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCA TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATCAT CCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCAG AGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACGTGGCAGTTTCTGAGCAGTCGTG CTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCGGCAGGTGG AGCTGTACCGCCTTCTYAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGGC GGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAG 

Gibbon

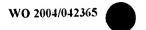
GTGGTTCCTAACATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA ATTGAAGATGGGGCATATCAGAGCCTAAGCCTCCTCTCCACCTTAATATTGACAGGAAAC CCCATCCAGAGTTTAGCTCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTAGTGG CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA AGAACTTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC ACTTGCAGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCT ATGAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGCCTTCRTAAGCTGACTTTAAGAA CCATCGTTTGGTTCTGGGAGAATTTAGAAATGAAGGAAACTTGGAAGAGTTTGACAAATC TGCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACCAC TACCTCGATGATATTATTGACTTATTTAATTGTTTGGCAAATGTTTCTTCATTTTCCCTGGT GAGTGTGACTATTAAAAGGGTAGAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAA TTAGTTAACTGTAAATTTGGACAGTTTCCCACATTGAACCTCAAATCTCTCAAAAGGCTTA CTTTCACTGCCAACAGAGTTGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTT TCTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGG ACAAACAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTATTACCATGAGTTCAAACT TCTTGGGCTTAGAACAGCTAGAACATCTGGATTTGCAGCATTCCAATTTGAAACAAATGA GTGAATTTCCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCAC ACCAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAATCTCGAAGTCTTGAAAATGG CTGGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGAC TCCAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTA TAAGTGTCTGAACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAA AAACAGGAACTACAGCGTTTTCCAAGTAGTCTAGCCTTCTTAAATCTTACTCAGAATGACT TTGCTTGTACTTGTGAACACGAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCAGCTCTT GGTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCT GAGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCAGTCCTCAGTGTG CTTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGC TGGCTGCATGAAGTATGGTAGAGGTGAAAACACCTATGATGCCTTTGTTATCTACTCCAGC CAGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCC CTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCYGGTGTGGCCATTGCTGCCAACATC ATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCC AGAGCCGCTGGTGTATCTTTGAGTATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCA TGCTGGGATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGT GGAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGATAGTGTCCTGGG GCGGCACATTTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCC AGAAGGAACAGTGGGTACAGGATGCAATTAG

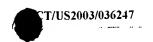




Gorilla

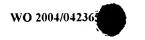
GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA ATTGAAGATGGGCCATATCAGAGCCTAAGCCACCTCCACCTTAATATTGACAGGAAAC CCCATCCAGAGTITAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA AGAACTTAATGTGGCTCACAATCTTATTCAATCTTTCAAATTACCTGAGTATTTTTCTAATC CTTGCGGGTTCTACATCAAATGCCCCTACTCAATCTCTTTTAGACCTGTCCCTGAACCCTA TGACCTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAA CGTCGTTTGGTTCTGGGAGAATTTAGAAATGAAGGAAACTTGGAAAAGTTTGACAAATCT GCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACTACT ACCTCGATGATATTATTGACTTATTTAATTGTTTGACAAATGTTTCTTCATTTTCCCTGGTG AGTGTGACTATTGAAAGGGTAAAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAAT TAGTTAACTGTAAATTTGGACAGTTTCCCACATTGAAACTCAAATCTCTCAAAAGGCTTAC TTTCACTTCCAACAAGGTGGGAATGCTTTTTCGGAAGTTGATCTACCAAGCCTTGAGTTT CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGGTGTTATTACCATGAGTTCAAACTT CTTGGGCTTAGAACAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC CAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAGTCTCGAAGTCTTGAAAATGGCT GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGACCT CAGTCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTAT AAGTGTCTGAACTCCCTCCGGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAA AACAGGAACTACAGCATTTTCCAAGCAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT TGCTTGTACTTGTGAACACCAGAGTTTCCTGCAATGGATCAAGGACCAGAGGCAGCTCTTG GTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCTG AGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCGGTCCTCAGTGTGC TTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT GGCTGCATAAAGTATGGTAGAGGTGAAAACGTCTATGATGCCTTTGTTATCTACTCAAGCC AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCA TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATCAT CCATGAAGGTTTCCATAAAAGTCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCAG AGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCGTG CTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGTGG AGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGGC GGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAG 





#### Rhesus monkey

GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACTTTAATATTGACAGGAAAC CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA AGAACTTAATGTGGCTCACAATCTTATCCAGTCTTTCAAATTACCTGAGTATTTTTCTAATC CTTGCAGGTTCTACATCAAATGCCCCTATCCAATCTCTCTTTAGACCTGTCCCTGAACCCTA TAAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAG CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAACTTGGAAGAGTTTGACAAATCT TCTCTGGAGGGATTGTGCAATTTGACCATTGAAGAATTCCGATTAACATACTTAGACTACT ACCTCGATAATATTATTGACTTATTTAATTGTTTGGCAAATGTTTCTTCATTTTCCCTGGTG AGTGTGAGTATTAAAAGGGTAGAAGACTTTTCTTATAATTTCAGATGGCAACATTTAGAAT TAGTTAACTGTAAATTTGAACAGTTTCCCACATTGGAACTCGAATCTCTCAAAAGGCTTAC TTTCACTGCCAACAAAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTATTACCATGAGTTCAAACTT CTTGGGCTTAGAAAACTAGAACATCTGGATTTCCAGCATTCCAATTTGAAACAGATGAG TCAATTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACA CCAGAGTTGCTTTCAATGGCATCTTCGATGGCTTGCTCAGTCTCAAAGTCTTAAAAATGGC TGGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGATCTGAAAAACTTGACC TTCCTGGACCTCTCTCAGTGTCAATTGGAGCAGTTGTCTCCAACAGCATTTGACACACTCA ACAAGCTTCAGGTACTAAATATGAGCCACAACAACTTCTTTTCATTGGATACGTTTCCTTA TAAGTGTCTGCCCTCCCAGGTTCTCGATTACAGTCTCAATCACATAATGACTTCCAAC AACCAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACT TTGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCAGCTCTT GGTGGAAGCTGAACGAATGGAATGTGCAACACCTTCAGATAAACAGGGCATGCCGGTGCT GAGTTTGAATATTACCTGTCAGATGAATAAGACCATCATTGGTGTGTCTGTTCAGTGTG CTTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGC TGGCTGCATAAASTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGC CAGGATGAGGACTGGGTAAGGAATGAACTAGTAAAGAATTTAGAAGAAGGGGTGCCTCC CTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCAAACATC ATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCC AGAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCG TGCAGGCATAATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCAGCAGGT GGAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGG GCAGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGTTGGATGGCAGATCGTGGAATCC **AGAAGAACAGTAG** 





....

#### Chimpanzee

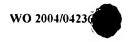
GTGGTTCCTAATATTACTTATCAATGCATGGAGCTGAATTTCTACAAAATCCCCGACAACC TCCCCTTCTCAACCAAGAACCTGGACCTGAGCTTTAATCCCCTGAGGCATTTAGGCAGCTA TAGCTTCTTCAGTTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGAAATCCAGACA ATTGAAGATGGGGCATATCAGAGCCTAAGCCACCTCTCCACCTTAATATTGACAGGAAAC CCCATCCAGAGTTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG CTGTGGAGACAAATCTAGCATCTCTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA AGAACTTAATGTGGCTCACAATCTTATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC CTTGCGGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA TGAACTTTATCCAACCAGGTGCATTTAAAGAAATTAGGCTTCATAAGCTGACTTTGAGAAA CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAGGAAACTTTGGAAAAGTTTTGACAAATCT GCTCTAGAGGGCCTGTGCAATTTGACCATTGAAGAATTCCGATTAGCATACTTAGACTACT ACCTCGATGATATTATTGACTTATTTAATTGTTTGACAAATGTTTCTTCATTTTCCCTGGTG AGTGTGACTATTAAAAGCGTAAAAGACTTTTCTTATAATTTCGGATGGCAACATTTAGAAT TAGTTAACTGTAAATTTGGACAGTTTCCCACATTGAAACTCAAATCTCTCAAAAGGCTTAC TTTCACTTCCAACAAGGTGGGAATGCTTTTTCAGAAGTTGATCTACCAAGCCTTGAGTTT CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAAAGTGATTTTGGGA CAACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGGTGTTATTACCATGAGTTCAAACTT CTTGGGCTTAGAACACTAGAACATCTGGA'ITTCCAGCATTCCAATTTGAAACAAATGAGT GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC CAGAGTTGCTTTCAATGGCATCTTCAATGGCTTGTCCAGTCTCGAAGTCTTGAAAATGGCT GGCAATTCTTTCCAGGAAAACTTCCTTCCAGATATCTTCACAGAGCTGAGAAACTTGACCT CAGTCTTCAGGTACTAAATATGAGCCACAACACTTCTTTTCATTGGATACGTTTCCTTAT AAGTGTCTGAACTCCCTCCAGGTTCTTGATTACAGTCTCAATCACATAATGACTTCCAAAA AACAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAGAATGACTT TGCTTGTACTTGTGAACACCAAAGTTTCCTGCAATGGATCAAGGACCAGAGGCAGCTCTTG GTGGAAGTTGAACGAATGGAATGTGCAACACCTTCAGATAAGCAGGGCATGCCTGTGCTG AGTTTGAATATCACCTGTCAGATGAATAAGACCATCATTGGTGTGTCGGTCCTCAGTGTGC TTGTAGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGCT GGCTGCATAAAGTATGGTAGAGGTGAAAACATCTATGATGCCTTTGTTATCTACTCAAGCC AGGATGAGGACTGGGTAAGGAATGAGCTAGTAAAGAATTTAGAAGAAGGGGTGCCTCCA TTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATCAT CCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTGTCCCAGCACTTCATCCAG AGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCGTG CTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGACCCTGCTCAGGCGGCAGGTGG AGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGGC GGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAAATCATGGAATCCAG 



T/US2003/036247

Capuchin

TGTGAAATCCACACAATTGAAGATGGTGCATATCAGAGCCTAAGCCACCTCTCCACCTTA ATATTGACAGGAAATCCTATCCAGAATTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTT TACAGAAACTGGTAGCTGTGGAGACACATCTGTTATCGCTAGAAAGCTTCCCCATTGGAC ATCTCAAAACTTTGAAGGACCTTAATGTGGCTCACAATCTAATCCAATCTTTCAAATTACC TGAGTATTTTCTAATCTGACCAATCTAGAGCACTTGGACCTTTCTAGTAACAATATTCAA AATATTTATTGCAAAGACTTGCAGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAG ACCTGTCCCTGAACCCTATAAACTTTATTCAGCCAGGTGCATTTAAAGAAATTAGGCTCCG TAAGCTGACTTTGAGAAATAATTTTGATAGTTTAAATGTAATGAAAACTTGCATTCACGGT GAAGACTTTGACAAATCTGCTCTGGAGGGCCTGTGCAATTTGACCATCAAAGAATTCCGA TTAGCATACTTAGACAACTTTCCAGATGATATTATTGACTTATTTAATTGTTTGGTAAATGT TTCTTCATTTTCCCTGTTGAGTGTGTATATTAAAAGAGTAGAAGACTTTTCTTATAATTTCA GATGGCAACATTTAGAATTAGTTAACTGTATATTTCAACAGTTTCCTCCACTGAAACTCAA ATCTCTCAAAAGGCTTACTTTCAGTAAAAACAAAGGTAGGAATCATTTTGCAGAAGTTGA TCTGCCAAGCCTTGAGTTTCTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGT TCTCAATCTGATTTTGGGACGACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGATGTTA TTACCATGAGTTCAAACTTCTTAGGCTTAGAACAACTAGAACACTTGGATTTCCAGCATTC CAATTTGAAACAAATGAGTGAGTTTTCAGTATTTCTATCACTCAGAAACCTCATTTACCTT GACATTTCTCATACTCACACCAGAGTTGCTTTCAATGGCATCTTTAATGGCTTGTTCAGTCT CAAAGTCTTGAAAATGGCTGGAAATTCTTTCCAGCAAAACTTCCTTGCAGATATCTTCACA GATCTGAATAACTTGATATTCCTGGACCTTTCTGAGTGTCAACTGGAGCAGTTGTCTCCAA CAGCATTTGACTCACTTCCCAGACTTCAGATACTAAATATGAGCCACAACAAGTTCTTTGC ATTGGATACATTTCCTTATAAGCATCTCTACTCCCTCCACGTTCTGGATTACAGTCTCAATC ACATAGGGACTTCCAAAAATCAGGAACTACAGCATTTTCCAAGTAGTCTAGCTTTCTTAAA TCTTACTCAAAATGACTTTGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAG GACCAGAGGCGGCTATTGGTGGAAGTTGAACGAATGGAATGCGCAACACCTTTAAATAGG AAGGGCATACCTGTGCTGAGTTTGAATATCACCTGTCAGATGAGTAAGACCATCATTGGT GTGTCAGTGCTCAGTGTGCTTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTT TCACCTGATGCTTCTTGCTGGCTGCATAAAGTATGGTAGAGGTGAAAACACCTATGATGCC TTTGTTATCTACTCAAGCCAGGATGAGGACTGGGTAAGGAATGAACTAGTAAAGAATTTA GAAGAAGGGGTGCCTCCTTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGG CCATTGCTGCCAACATCATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGT ATCCCAGCACTTCATCCAGAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGG CAGTTTCTGAGCAGTCGTGCTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGTCCC AGGACAGTGTCCTGGGGGGGCATATCTTCTGGAGGCGACTCAGAAAAGCCCTGCTGAATG GTAGACCGTGGAGTCCAGAAGGAACAGTGGGTGCAGGATGCGATTAG





Squirrel monkey

GTGGTTCCTAACGTTACTTATCAATGCATGGAACTGAATYTCTACAAAATCCCCGACAACA TCCCCTTCTCAACTAAGAACCTGGACCTGAGCTTTAACCCCCTGAGGCATTTAGGCAGCCA TAGCTTCTTCAATTTCCCAGAACTGCAGGTGCTGGATTTATCCAGGTGTGACATCCAGACA ATCGAAGATGGGGCATATCAGAGCCTAAGCCACCTCCACCTTAATATTGACAGGAAAT CCTATCCAGAATTTAGCCCTGGGAGCCTTTTCTGGACTATCAAGTTTACAGAAGCTGGTGG CTGTGGAGACACATCTGTTATCACTAGAGAACTTCCCCATTGGACATCTCAAAACTTTGAA GGACCTTAATGTGGCTCACAATCTAATCCAATCTTTCAAATTACCTGAGTATTTTTCTAATC CTTGCAGGTTCTACATCAAATGCCCCTACTCAATCTCTCTTTAGACCTGTCCCTGAACCCTA TAAACTTTATTCAACCAGGTGCGTTTAAAGAAATTAGGCTCCATAAGCTGACTTTGAGAAA CATCGTTTGGTTCTGGGAGAATTTAGAAATGAAAGAAATATTGAAGACTTTGACAAATCT TTCTAGATGATATTATTGACTTATTTAACTGTTTAGCAAATGTTTCTTCATTTTCCCTGGTG AATGTGCATATTAAAAGAGTAGAAGACTTTTCTTATAATTTTAGATGGCAACATTTAGAAT TAGTTAACTGTGTATTTCAACAGTTTCCTCCACTGAAACTCAAATCTCTCAAAAGGCTTAC TTTCACTGCCAACAAGGTAGGAATCATTTTTCAGAAGTTGATCTTCCAAGCCTTGAGTTT CTAGATCTCAGTAGAAATGGCTTGAGTTTCAAAGGTTGCTGTTCTCAATCTGATTTTGGGA CGACCAGCCTAAAGTATTTAGATCTGAGCTTCAATGACGTTATTACCATGGGTTCAAACTT CTTAGGCTTAGAACAACTAGAACACTTGGATTTCCAGCATTCCAATTTGAAACAAATGAGT GAGTTTTCAGTATTCCTATCACTCAGAAACCTCATTTACCTTGACATTTCTCATACTCACAC CAGAGTTGCTTTCAATGGCATCTTTAATGGCTTGTTCAGTCTCAAAGTCTTGAAAATGGCT GGAAATTCTTTCCAGCAAAACTTCCTTGAAGATATCTTCACRGATCTGAATAACTTGATAT TCCTGGACCTCTCTGAGTGTCAGCTGGAGCAGTTGTCTCCAACAGCATTTGACTCACTTCC CAGACTTCGGATACTAAATATGAGCCACAACAACTTCTTTGCATTGGATACATTCCCTTAC AAGCATCTCTACTCCCTCCAGGTTCTGGATTACAGTCTCAATCATATAGGGACTTCCAAAA ATCAGGAACTGCAGCATTTTCCAAGTAGTCTAGCTTTCTTAAATCTTACTCAAAATGACTT TGCTTGTACTTGTGAACACCAGAGTTTCCTGCAGTGGATCAAGGACCAGAGGCGGCTGTT GGTGGAAGTTGAACAAATGGAATGTGCAACACCTTTAAATAGGAAGGCCATACCTGTGCT GAGTTTGAATATCACCTGTCAGATGAGTAAGACTATCATTGGTGTGTCAGTGCTCAGTGTG CTTGTGGTATCTGTTGTAGCAGTTCTGGTCTATAAGTTCTATTTTCACCTGATGCTTCTTGC TGGCTGCATAAAGTATGGTAGAGGTGAAAACACCTATGATGCCTTTGTTATCTACTCAAGC CAGGATGAGGACTGGGTAAGGAATGAACTAGTAAAGAATTTAGAAGAAGGGGTGCCTCC CTTTCAGCTCTGCCTTCACTACAGAGACTTTATTCCCGGTGTGGCCATTGCTGCCAACATC ATCCATGAAGGTTTCCATAAAAGCCGAAAGGTGATTGTTGTGGTATCTCAGCACTTCATCC AGAGCCGCTGGTGTATCTTTGAATATGAGATTGCTCAGACCTGGCAGTTTCTGAGCAGTCG TGCTGGTATCATCTTCATTGTCCTGCAGAAGGTGGAGAAGTCCCTGCTCAGGCAGCAGGTG GAGCTGTACCGCCTTCTCAGCAGGAACACTTACCTGGAGTGGGAGGACAGTGTCCTGGGG AGGCACATCTTCTGGAGACGACTCAGAAAAGCCCTGCTGGATGGTAGACCGTGGAATCCA GAAGGAACAGTGGGTGCAGGATGCGAATAG

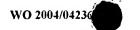
FIGURE 9

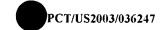
10/532153 JC13 Rec'd PCT/PTO 20 APR 2005

## SEQUENCE LISTING

WO 2004/042365

. 4 4 0 :	D 3	. Li annou. Can	omics IIC										
<110>	c mbassassing for the Treatment Of												
<120>	Devel Endot	lopment of T coxin-Mediat	Therapeutics ted Diseases	s for the Tr	reatment of								
<130>	GENO	200.3.1											
<150> <151>		00,422 -03-18											
<160>	40												
<170>	Pate	ntIn versio	n 3.2										
<210> <211> <212> <213>	1 2427 DNA Pan	troglodytes											
<400> gtggtt	1 ccta	atattactta	tcaatgcatg	gagctgaatt	tctacaaaat	ccccgacaac	60						
		caaccaagaa					120						
		tcagtttccc					180						
		atggggcata					240						
		agagtttagc					300						
		agacaaatct					360						
ttgaaa	agaac	ttaatgtggc	tcacaatctt	atccaatctt	tcaaattacc	tgagtatttt	420						
tctaat	tctga	ccaatctaga	gcacttggac	ctttccagca	acaagattca	aagtatttat	480						
tgcaca	agact	tgcgggttct	acatcaaatg	cccctactca	atctctcttt	agacctgtcc	540						
ctgaa	cccta	tgaactttat	ccaaccaggt	gcatttaaag	aaattaggct	tcataagctg	600						
acttt	gagaa	ataattttga	tagtttaaat	gtaatgaaaa	cttgtattca	aggtctggct	660						
ggttt	agaag	tccatcgttt	ggttctggga	gaatttagaa	atgaaggaaa	cttggaaaag	720						
tttga	caaat	ctgctctaga	gggcctgtgc	aatttgacca	ttgaagaatt	ccgattagca	780						
tactt	agact	actacctcga	tgatattatt	gacttattta	attgtttgac	aaatgtttct	840						
tcatt	ttccc	tggtgagtgt	gactattaaa	agcgtaaaag	acttttctta	taatttcgga	900						
tggca	acatt	tagaattagt	taactgtaaa	tttggacagt	ttcccacatt	gaaactcaaa	960						
tctct	caaaa	ggcttacttt	cacttccaac	aaaggtggga	atgctttttc	agaagttgat	1020						
ctacc	aagcc	ttgagtttct	agatctcagt	agaaatggct	tgagtttcaa	aggttgctgt	1080						
tctca	aagtg	attttgggac	aaccagccta	aagtatttag	atctgagctt	caatggtgtt	1140						





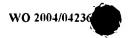
attaccatga gttcaaactt	cttgggctta	gaacaactag	aacatctgga	tttccagcat	1200
tccaatttga aacaaatgag	tgagttttca	gtattcctat	cactcagaaa	cctcatttac	1260
cttgacattt ctcatactca	caccagagtt	gctttcaatg	gcatcttcaa	tggcttgtcc	1320
agtctcgaag tcttgaaaat	ggctggcaat	tctttccagg	aaaacttcct	tccagatatc	1380
ttcacagage tgagaaactt	gaccttcctg	gacctctctc	agtgtcaact	ggagcagttg	1440
tctccaacag catttaactc	actctccagt	cttcaggtac	taaatatgag	ccacaacaac	1500
ttcttttcat tggatacgtt	tccttataag	tgtctgaact	ccctccaggt	tcttgattac	1560
agtotcaato acataatgac	ttccaaaaaa	caggaactac	agcattttcc	aagtagtcta	1620
gctttcttaa atcttactca	gaatgacttt	gcttgtactt	gtgaacacca	aagtttcctg	1680
caatggatca aggaccagag	gcagctcttg	gtggaagttg	aacgaatgga	atgtgcaaca	1740
ccttcagata agcagggcat	acctatacta	agtttgaata	tcacctgtca	gatgaataag	1800
					1860
accatcattg gtgtgtcggt					
tataagttct attttcacct	gatgcttctt	gctggctgca	taaagtatgg	tagaggtgaa	1920
aacatctatg atgcctttgt	tatctactca	agccaggatg	aggactgggt	aaggaatgag	1980
ctagtaaaga atttagaaga	aggggtgcct	ccatttcagc	tctgccttca	ctacagagac	2040
tttattcccg gtgtggccat	tgctgccaac	atcatccatg	aaggtttcca	taaaagccga	2100
aaggtgattg ttgtggtgtc	ccagcacttc	atccagagcc	gctggtgtat	ctttgaatat	2160
gagattgctc agacctggca	gtttctgagc	agtcgtgctg	gtatcatctt	cattgtcctg	2220
cagaaggtgg agaagaccct	gctcaggcgg	caggtggagc	tgtaccgcct	tctcagcagg	2280
aacacttacc tggagtggga	ggacagtgtc	ctggggcggc	acatcttctg	gagacgactc	2340
agaaaagccc tgctggatgg	taaatcatgg	aatccagaag	gaacagtggg	tacaggatgc	2400
aattggcagg aagcaacatc	tatctga				2427

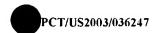
<210> 2 <211> 2427 <212> DNA <213> Pan troglodytes

<220> <221> CDS <222> (1)..(2427)

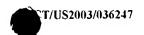
<400> 2 gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa 48 Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys

1		10		1.5
1	5	10		15
atc ccc gac aac Ile Pro Asp Asn 20	ctc ccc ttc Leu Pro Phe	tca acc aag Ser Thr Lys 25	aac ctg gac ctg Asn Leu Asp Leu 30	agc ttt 96 Ser Phe
aat ccc ctg agg Asn Pro Leu Arg 35	cat tta ggc His Leu Gly	agc tat agc Ser Tyr Ser 40	ttc ttc agt ttc Phe Phe Ser Phe 45	cca gaa 144 Pro Glu
ctg cag gtg ctg Leu Gln Val Leu 50	gat tta tcc Asp Leu Ser 55	agg tgt gaa Arg Cys Glu	atc cag aca att Ile Gln Thr Ile 60	gaa gat 192 Glu Asp
ggg gca tat cag Gly Ala Tyr Gln 65	agc cta agc Ser Leu Ser 70	cac ctc tcc His Leu Ser	acc tta ata ttg Thr Leu Ile Leu 75	aca gga 240 Thr Gly 80
aac ccc atc cag Asn Pro Ile Gln	agt tta gcc Ser Leu Ala 85	ctg gga gcc Leu Gly Ala 90	ttt tct gga cta Phe Ser Gly Leu	tca agt 288 Ser Ser 95
tta cag aag ctg Leu Gln Lys Leu 100	gtg gct gtg Val Ala Val	gag aca aat Glu Thr Asn 105	cta gca tct cta Leu Ala Ser Leu 110	gag aac 336 Glu Asn
ttc ccc att gga Phe Pro Ile Gly 115	His Leu Lys	act ttg aaa Thr Leu Lys 120	gaa ctt aat gtg o Glu Leu Asn Val 7 125	gct cac 384 Ala His
aat ctt atc caa Asn Leu Ile Gln 130	tct ttc aaa Ser Phe Lys 135	tta cct gag Leu Pro Glu	tat ttt tct aat o Tyr Phe Ser Asn 1 140	ctg acc 432 Leu Thr
aat cta gag cac Asn Leu Glu His 145	ttg gac ctt : Leu Asp Leu : 150	Ser Ser Asn	aag att caa agt a Lys Ile Gln Ser 1 155	att tat 480 Nie Tyr 160
tgc aca gac ttg Cys Thr Asp Leu	cgg gtt cta o Arg Val Leu I 165	cat caa atg His Gln Met 170	ccc cta ctc aat o Pro Leu Leu Asn 1	etc tct 528 Leu Ser .75
tta gac ctg tcc Leu Asp Leu Ser 180	ctg aac cct a Leu Asn Pro N	atg aac ttt a Met Asn Phe 1 185	atc caa cca ggt o Ile Gln Pro Gly <i>F</i> 190	gca ttt 576 Ala Phe
aaa gaa att agg Lys Glu Ile Arg 195	Leu His Lys I	ctg act ttg a Leu Thr Leu <i>i</i> 200	aga aat aat ttt o Arg Asn Asn Phe A 205	gat agt 624 usp Ser
tta aat gta atg Leu Asn Val Met 210	aaa act tgt a Lys Thr Cys 1 215	att caa ggt ( [le Gln Gly ]	ctg gct ggt tta g Leu Ala Gly Leu G 220	aa gtc 672 ilu Val
cat cgt ttg gtt His Arg Leu Val 225	ctg gga gaa t Leu Gly Glu E 230	Phe Arg Asn (	gaa gga aac ttg g Glu Gly Asn Leu G 235	aa aag 720 lu Lys 240
ttt gac aaa tct Phe Asp Lys Ser	gct cta gag g Ala Leu Glu G	ggc ctg tgc a Gly Leu Cys A	aat ttg acc att g Asn Leu Thr Ile G	aa gaa 768 lu Glu

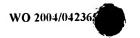


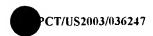


250 255 ttc cga tta gca tac tta gac tac tac ctc gat gat att att gac tta 816 Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu 260 265 ttt aat tgt ttg aca aat gtt tct tca ttt tcc ctg gtg agt gtg act 864 Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr 280 att aaa agc gta aaa gac ttt tct tat aat ttc gga tgg caa cat tta 912 Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu 300 gaa tta gtt aac tgt aaa ttt gga cag ttt ccc aca ttg aaa ctc aaa 960 Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys 310 tot oto aaa agg ott act tic act too aac aaa ggt ggg aat got tit 1008 Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe 330 tca gaa gtt gat cta cca agc ctt gag ttt cta gat ctc agt aga aat 1056 Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 ggc ttg agt ttc aaa ggt tgc tgt tct caa agt gat ttt ggg aca acc 1104 Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr 360 age cta aag tat tta gat etg age tte aat ggt gtt att ace atg agt 1152 Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser 375 tca aac ttc ttg ggc tta gaa caa cta gaa cat ctg gat ttc cag cat 1200 Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His 390 395 tcc aat ttg aaa caa atg agt gag ttt tca gta ttc cta tca ctc aga 1248 Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg 405 410 aac ctc att tac ctt gac att tct cat act cac acc aga gtt gct ttc 1296 Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe 420 425 aat ggc atc ttc aat ggc ttg tcc agt ctc gaa gtc ttg aaa atg gct 1344 Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala 435 440 ggc aat tot tto cag gaa aac tto ott coa gat ato tto aca gag otg 1392 Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu 450 455 aga aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag cag ttg 1440 Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu tot oca aca goa ttt aac toa oto too agt ott cag gta ota aat atg 1488 Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met



				485					490					495		,	
agc c Ser H																	1536
aac t Asn S	er																1584
aaa a Lys L 5																	1632
ctt a Leu T 545																	1680
caa t Gln T	gg Tp	atc Ile	aag Lys	gac Asp 565	cag Gln	agg Arg	cag Gln	ctc Leu	ttg Leu 570	gtg Val	gaa Glu	gtt Val	gaa Glu	cga Arg 575	atg Met		1728
gaa t Glu C																	1776
aat a Asn I	tc [le	acc Thr 595	tgt Cys	cag Gln	atg Met	aat Asn	aag Lys 600	acc Thr	atc Ile	att Ile	ggt Gly	gtg Val 605	tcg Ser	gtc Val	ctc Leu		1824
agt g Ser V 6	gtg /al 510	ctt Leu	gta Val	gta Val	tct Ser	gtt Val 615	gta Val	gca Ala	gtt Val	ctg Leu	gtc Val 620	tat Tyr	aag Lys	ttc Phe	tat Tyr		1872
ttt c Phe H 625																	1920
aac a Asn I																	1968
gta a Val A	agg Arg	aat Asn	gag Glu 660	cta Leu	gta Val	aag Lys	aat Asn	tta Leu 665	gaa Glu	gaa Glu	ggg ggg	gtg Val	cct Pro 670	cca Pro	ttt Phe		2016
cag c Gln I																	2064
gcc a Ala A					-					_	_						2112
gtg g Val V 705																	2160
gag a Glu I																	2208





725	730	735

														cag Gln		2256
														gag Glu		2304
														gcc Ala		2352
ctg Leu 785	gat Asp	ggt Gly	aaa Lys	tca Ser	tgg Trp 790	aat Asn	cca Pro	gaa Glu	gga Gly	aca Thr 795	gtg Val	ggt Gly	aca Thr	gga Gly	tgc Cys 800	2400
				gca Ala 805				tga								2427

<210> 3

<211> 808

<212> PRT

<213> Pan troglodytes

<400> 3

Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1 5 10 15

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 20 25 30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu 35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp 50 55 60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly 65 70 75 80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser 85 90 95

Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn 100 105 110

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His

115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr 130 135 140

- Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr 145 150 155 160
- Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser 165 170 175
- Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe 180 185 190
- Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser 195 200 205
- Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val 210 215 220
- His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys 225 230 235 240
- Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu 245 250 255
- Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu 260 265 270
- Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr 275 280 285
- Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu 290 295 300
- Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys 305 310 315 320
- Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe 325 330 335
- Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 350
- Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr

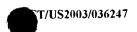
355

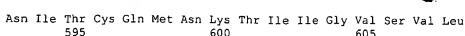
360

365

Ser	Leu	Lys	Tyr	Leu	Asp	Leu	Ser	Phe	Asn	Gly	Val	Ile	Thr	Met	Ser
	370					375				-	380				

- Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His 385 390 395 400
- Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg 405 410 415
- Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe 420 425 430
- Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala 435 440 445
- Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu 450 455 460
- Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 470 475 480
- Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met 485 490 495
- Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu 500 505 510
- Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 515 520 525
- Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn 530 535 540
- Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu 545 550 555 560
- Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met 565 570 575
- Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu 580 585 590





Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr 610 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu 625 630 635

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp 645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe 660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala 675 680 685

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Val 740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp
755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 775 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys 785 790 795 800

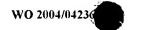
Asn Trp Gln Glu Ala Thr Ser Ile 805

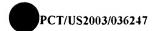
<210> 4

<211> 2427

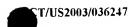
<212> DNA

<213> Gorilla gorilla

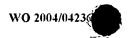


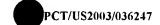


gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60 ctcccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120 180 tatagettet teagttteee agaactgeag gtgetggatt tateeaggtg tgaaateeag acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240 300 aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg qtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360 ttgaaagaac ttaatgtggc tcacaatctt attcaatctt tcaaattacc tgagtatttt 420 tctaatctga ccaatctaga gtacttggac ctttccagca acaagattca aagtatttat 480 tgcacagact tgcgggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540 ctqaacccta tgacctttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600 actttgagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660 ggtttagaag tccgtcgttt ggttctggga gaatttagaa atgaaggaaa cttggaaaag 720 780 tttgacaaat ctgctctaga gggcctgtgc aatttgacca ttgaagaatt ccgattagca tacttagact actacctcga tgatattatt gacttattta attgtttgac aaatgtttct 840 900 tcattttccc tggtgagtgt gactattgaa agggtaaaag acttttctta taatttcgga tqqcaacatt taqaattaqt taactqtaaa tttqqacaqt ttcccacatt qaaactcaaa 960 tototoaaaa qqottaottt cacttooaac aaaggtggga atgottttto ggaagttgat 1020 1080 ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt totcaaaqtg attttgggac aaccagoota aagtatttag atotgagott caatggtgtt 1140 1200 attaccatga gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat tccaatttga aacaaatgag tgagttttca gtattcctat cactcagaaa cctcatttac 1260 1320 cttqacattt ctcatactca caccagagtt gctttcaatg gcatcttcaa tggcttgtcc agtotogaag tottgaaaat ggotggoaat totttocagg aaaacttoot tocagatato 1380 ttcacagagc tgagaaactt gaccttcctg gacctctctc agtgtcaact ggagcagttg 1440 totocaacag catttaactc actotocagt ottoaggtac taaatatgag coacaacaac 1500 ttetttteat tqqatacqtt teettataaq tqtetqaact ceeteegggt tettgattae 1560 agtotoaato acataatgao ttooaaaaaa caggaactao agcattttoo aagcagtota 1620 gctttcttaa atcttactca gaatgacttt gcttgtactt gtgaacacca gagtttcctg 1680 1740 caatggatca aggaccagag gcagctcttg gtggaagttg aacgaatgga atgtgcaaca

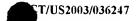


	1800
accatcattg gtgtgtcggt cctcagtgtg cttgtagtat ctgttgtagc agttctggtc	1860
tataagttot attttcacct gatgottott gotggotgoa taaagtatgg tagaggtgaa	1920
aacgtctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgag	1980
ctagtaaaga atttagaaga aggggtgcct ccatttcagc tctgccttca ctacagagac	2040
tttattcccg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagtcga	2100
aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat	2160
gagattgctc agacctggca gtttctgagc agtcgtgctg gtatcatctt cattgtcctg	2220
cagaaggtgg agaagaccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg	2280
aacacttacc tggagtggga ggacagtgtc ctggggcggc acatcttctg gagacgactc	2340
agaaaagccc tgctggatgg taaatcatgg aatccagaag gaacagtggg tacaggatgc	2400
aattggcagg aagcaacatc tatctga	2427
<210> 5 <211> 2427 <212> DNA <213> Gorilla gorilla	
<221> CDS <222> (1)(2427)	
<221> CDS	48
<pre>&lt;221&gt; CDS &lt;222&gt; (1)(2427)  &lt;400&gt; 5 gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys</pre>	48 96
<pre>&lt;221&gt; CDS &lt;222&gt; (1)(2427)  &lt;400&gt; 5 gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1</pre>	
<pre>&lt;221&gt; CDS &lt;222&gt; (1)(2427)  &lt;400&gt; 5 gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1</pre>	96
<pre>&lt;221&gt; CDS &lt;222&gt; (1)(2427)  &lt;400&gt; 5 gtg gtt cct aat att act tat caa tgc atg gag ctg aat ttc tac aaa Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1</pre>	96 144

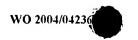


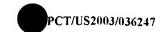


tta Lei	a caq ı Glr	g aaq n Lys	g cto s Leu 100	ı Val	g gct . Ala	gto Val	gag Glu	aca Thi	r Asn	cta Leu	a gca a Ala	a to: a Se:	ct. Lei	ı Gl	g aac 1 Asn	336
tto Phe	e ccc	2 att	e GT?	cat His	cto Leu	aaa Lys	act Thr 120	Lei	g aaa 1 Lys	gaa Glu	ctt Leu	aat 1 Asr 125	ı Va.	g gci L Ala	cac His	384
aat Asn	t ctt Leu 130	Ile	caa Gln	tct Ser	tto Phe	aaa Lys 135	Leu	cct Pro	gag Glu	tat Tyr	ttt Phe	Ser	: aat : Asr	cto Lei	g acc 1 Thr	432
aat Asn 145	Leu	gag Glu	tac Tyr	ttg Leu	gac Asp 150	Leu	tcc Ser	ago Ser	aac Asn	aag Lys 155	Ile	caa Gln	agt Ser	: att	tat Tyr 160	480
tgc Cys	aca Thr	gac Asp	ttg Leu	cgg Arg 165	gtt Val	cta Leu	cat His	caa Gln	atg Met 170	ccc Pro	cta Leu	cto Leu	aat Asn	teto Leu 175	tct Ser	528
tta Leu	gac Asp	ctg Leu	tcc Ser 180	Leu	aac Asn	cct Pro	atg Met	acc Thr 185	Phe	atc Ile	caa Gln	cca Pro	ggt Gly 190	Ala	ttt Phe	576
aaa Lys	gaa Glu	att Ile 195	agg Arg	ctt Leu	cat His	aag Lys	ctg Leu 200	act Thr	ttg Leu	aga Arg	aat Asn	aat Asn 205	ttt Phe	gat Asp	agt Ser	624
tta Leu	aat Asn 210	gta Val	atg Met	aaa Lys	act Thr	tgt Cys 215	att Ile	caa Gln	ggt Gly	ctg Leu	gct Ala 220	ggt Gly	tta Leu	gaa Glu	gtc Val	672
Arg 225	Arg	Leu	Val	Leu	Gly 230	Glu	Phe	Arg	aat Asn	G1u 235	Gly	Asn	Leu	Glu	Lys 240	720
ttt Phe	gac Asp	aaa Lys	tct Ser	gct Ala 245	cta Leu	gag Glu	ggc Gly	ctg Leu	tgc Cys 250	aat Asn	ttg Leu	acc Thr	att Ile	gaa Glu 255	gaa Glu	768
ttc Phe	cga Arg	tta Leu	gca Ala 260	Tyr	Leu	Asp	tac Tyr	Tyr	ctc Leu	gat Asp	gat Asp	Ile	att Ile 270	Asp	tta Leu	816
ttt Phe	aat Asn	tgt Cys 275	ttg Leu	aca Thr	aat Asn	gtt Val	tct Ser 280	tca Ser	ttt Phe	tcc Ser	ctg Leu	gtg Val 285	agt Ser	gtg Val	act Thr	864
lle	G1u 290	Arg	Val	Lys	Asp	Phe 295	Ser	Tyr	aat Asn	Phe	Gly 300	Trp	Gln	His	Leu	912
305	Leu	vaı	Asn	Cys	Lys 310	Phe	Gly	Gln		Pro 315	Thr	Leu	Lys	Leu	Lys 320	960
tct Ser	ctc Leu	aaa Lys	Arg	ctt Leu 325	act Thr	ttc Phe	act Thr	tcc Ser	aac Asn 330	aaa Lys	ggt Gly	ggg Gly	aat Asn	gct Ala 335	ttt Phe	1008



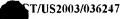
tco Ser	g gaa Glu	a gti u Val	gat L Asp 340	) Le	a cca ı Pro	a ago Ser	ctt Leu	gaq Glu 345	Phe د	cta Lev	a gat ı Ası	t cto	agi Sei 350	r Ar	a aat g Asn	1056
ggc Gly	tto Leu	g agt sei 355	r Phe	aaa Lys	ggt Gly	tgc Cys	tgt Cys 360	Sei	caa Gln	agt Ser	gat Asp	ttt Phe 365	Gly	g aca / Thi	a acc Thr	1104
ago Ser	cta Leu 370	ı Lys	g tat s Tyr	tta Leu	gat Asp	ctg Leu 375	Ser	tto Phe	aat Asn	ggt Gly	gtt Val 380	. Ile	aco Thi	ato Met	g agt : Ser	1152
tca Ser 385	Asn	tto Phe	ttg Leu	ggc Gly	tta Leu 390	Glu	caa Gln	cta Leu	gaa Glu	cat His	Leu	gat Asp	tto Phe	caç Glr	cat His 400	1200
tcc Ser	aat Asn	ttg Lev	aaa Lys	Gln 405	Met	agt Ser	gag Glu	t t t Phe	tca Ser 410	Val	ttc Phe	cta Leu	tca Ser	cto Leu 415	aga Arg	1248
aac Asn	ctc Leu	att	tac Tyr 420	Leu	gac Asp	att Ile	tct Ser	cat His 425	Thr	cac His	acc Thr	aga Arg	gtt Val 430	Ala	ttc Phe	1296
aat Asn	ggc Gly	atc Ile 435	Phe	aat Asn	ggc Gly	ttg Leu	tcc Ser 440	agt Ser	ctc Leu	gaa Glu	gtc Val	ttg Leu 445	aaa Lys	atg Met	gct Ala	1344
ggc Gly	aat Asn 450	tct Ser	ttc Phe	cag Gln	gaa Glu	aac Asn 455	ttc Phe	ctt Leu	cca Pro	gat Asp	atc Ile 460	Phe	aca Thr	gag Glu	ctg Leu	1392
aga Arg 465	aac Asn	ttg Leu	acc Thr	ttc Phe	ctg Leu 470	gac Asp	ctc Leu	tct Ser	cag Gln	tgt Cys 475	caa Gln	ctg Leu	gag Glu	cag Gln	ttg Leu 480	1440
tct Ser	cca Pro	aca Thr	gca Ala	ttt Phe 485	aac Asn	tca Ser	ctc Leu	tcc Ser	agt Ser 490	ctt Leu	cag Gln	gta Val	cta Leu	aat Asn 495	atg Met	1488
agc Ser	cac His	aac Asn	aac Asn 500	ttc Phe	ttt Phe	tca Ser	ttg Leu	gat Asp 505	Thr	ttt Phe	cct Pro	tat Tyr	aag Lys 510	Cys	ctg Leu	1536
aac Asn	tcc Ser	ctc Leu 515	cgg Arg	gtt Val	ctt Leu	gat Asp	tac Tyr 520	agt Ser	ctc Leu	aat Asn	cac His	ata Ile 525	atg Met	act Thr	tcc Ser	1584
aaa Lys	aaa Lys 530	cag Gln	gaa Glu	cta Leu	cag Gln	cat His 535	ttt Phe	cca Pro	agc Ser	agt Ser	cta Leu 540	gct Ala	ttc Phe	tta Leu	aat Asn	1632
ctt Leu 545	act Thr	cag Gln	aat Asn	gac Asp	ttt Phe 550	gct Ala	tgt Cys	act Thr	tgt Cys	gaa Glu 555	cac His	cag Gln	agt Ser	ttc Phe	ctg Leu 560	1680
caa Gln	tgg Trp	atc Ile	aag Lys	gac Asp 565	cag Gln	agg Arg	cag Gln	ctc Leu	ttg Leu 570	gtg Val	gaa Glu	gtt Val	gaa Glu	cga Arg 575	atg Met	1728





 $a \gtrsim_{\sigma}$ 

						gat Asp										1776
			_	_	-	aat Asn	_						-	-		1824
						gtt Val 615										1872
		_	_			gct Ala		_		_			_		-	1920
						gtt Val										1968
_					-	aag Lys			_	•						2016
						aga Arg										2064
_					-	ggt Gly 695				-	-	_			-	2112
			-			atc Ile	-	_	-		_			-		2160
		-	-			cag Gln		_	-	_	-	-				2208
		_	_	-	_	gtg Val		_		-			_	-		2256
	-					agc Ser										2304
_	-	-				atc Ile 775			-	_		-		-	-	2352
_	•					aat Asn		-			-				_	2400
		-	_	-		tct Ser		tga								2427

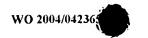


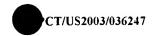
														_	
<21: <21: <21: <21:	1> 2>	6 808 PRT Gori	lla (	gori	lla										
<40	0>	6													
Val 1	Val	Pro	Asn	Ile 5	Thr	Tyr	Gln	Cys	Met 10	Glu	Leu	Asn	Phe	Tyr 15	Lys
Ile	Pro	Asp	Asn 20	Leu	Pro	Phe	Ser	Thr 25	Lys	Asn	Leu	Asp	Leu 30	Ser	Phe
Asn	Pro	Leu 35	Arg	His	Leu	Gly	Ser 40	Туr	Ser	Phe	Phe	Ser 45	Phe	Pro	Glu
Leu	Gln 50	Val	Leu	Asp	Leu	Ser 55	Arg	Cys	Glu	Ile	Gln 60	Thr	Ile	Glu	Asp
Gly 65	Ala	Tyr	Gln	Ser	Leu 70	Ser	His	Leu	Ser	Thr 75	Leu	Ile	Leu	Thr	Gly 80
Asn	Pro	Ile	Gln	Ser 85	Leu	Ala	Leu	Gly	Ala 90	Phe	Ser	Gly	Leu	Ser 95	Ser
Leu	Gln	Lys	Leu 100	Val	Ala	Val	Glu	Thr 105	Asn	Leu	Ala	Ser	Leu 110	Glu	Asn
Phe	Pro	Ile 115	Gly	His	Leu	Lys	Thr 120	Leu	Lys	Glu	Leu	Asn 125	Val	Ala	His
Asn	Leu 130	Ile	Gln	Ser	Phe	Lys 135	Leu	Pro	Glu	Tyr	Phe 140	Ser	Asn	Leu	Thr
Asn 145	Leu	Glu	Tyr	Leu	Asp 150	Leu	Ser	Ser	Asn	Lys 155	Ile	Gln	Ser	Ile	Tyr 160
Cys	Thr	Asp	Leu	Arg 165	Val	Leu	His	Gln	Met 170	Pro	Leu	Leu	Asn	Leu 175	Ser
Leu	Asp	Leu	Ser 180	Leu	Asn	Pro	Met	Thr 185	Phe	Ile	Gln	Pro	Gly 190	Ala	Phe

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser

200

195





Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val 210 215 220

Arg Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys 225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu 245 250 255

Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu 260 265 270

Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr 275 280 285

Ile Glu Arg Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu
290 295 300

Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys 305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe 325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr 355 360 365

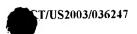
Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser 370 375 380

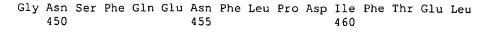
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His 385 390 395 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg 405 410 415

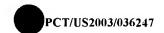
Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe 420 425 430

Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala 435 440 445





- Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 470 475 480
- Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met 485 490 495
- Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu 500 505 510
- Asn Ser Leu Arg Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 515 520 525
- Lys Lys Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn 530 540
- Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu 545 550 555 560
- Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met 565 570 575
- Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu 580 585
- Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu 595 600 605
- Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr 610 620
- Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu 630 635 640
- Asn Val Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp 645 650 655
- Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe 660 665 670
- Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala 675 680 685



Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val 690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile
725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val 740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys 785 790 795 800

Asn Trp Gln Glu Ala Thr Ser Ile 805

<210> 7

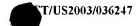
<211> 2406

<212> DNA

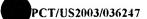
<213> Hylobates lar

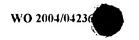
<400> 7

gtggttccta acattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60 ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120 tatagettet teagttteee agaactgeag gtgetggatt tateeaggtg tgaaateeag 180 acaattgaag atggggcata tcagagccta agcctcctct ccaccttaat attgacagga 240 aaccccatcc agagtttagc tctgggagcc ttttctggac tatcaagttt acagaagcta 300 gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360 ttgaaagaac ttaatgtggc tcacaatctt atccaatctt tcaaattacc tgagtatttt 420 tctaatctga ccaatctaga gcacttggac ctttccagca acaagattca aagtatttat 480 tgcaaagact tgcaggttct acatcaaatg cccctactca atctctcttt agacctgtcc 540 ctgaacccta tgaactttat ccaaccaggt gcatttaaag aaattagcct tcrtaagctg 600 actttaagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660

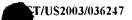


ggtttagaag	tccatcgttt	ggttctggga	gaatttagaa	atgaaggaaa	cttggaagag	720
tttgacaaat	ctgctctaga	gggcctgtgc	aatttgacca	ttgaagaatt	ccgattagca	780
tacttagacc	actacctcga	tgatattatt	gacttattta	attgtttggc	aaatgtttct	840
tcattttccc	tggtgagtgt	gactattaaa	agggtagaag	acttttctta	taatttcgga	900
tggcaacatt	tagaattagt	taactgtaaa	tttggacagt	ttcccacatt	gaacctcaaa	960
tctctcaaaa	ggcttacttt	cactgccaac	agaggtggga	atgctttttc	agaagttgat	1020
ctaccaagcc	ttgagtttct	agatctcagt	agaaatggct	tgagtttcaa	aggttgctgt	1080
tctcaaagtg	attttgggac	aaacagccta	aagtatttag	atctgagctt	caatgatgtt	1140
attaccatga	gttcaaactt	cttgggctta	gaacagctag	aacatctgga	tttgcagcat	1200
tccaatttga	aacaaatgag	tgaattttca	gtattcctat	cactcagaaa	cctcatttac	1260
cttgacattt	ctcatactca	caccagagtt	gctttcaatg	gcatcttcaa	tggcttgtcc	1320
aatctcgaag	tcttgaaaat	ggctggcaat	tctttccagg	aaaacttcct	tccagatatc	1380
ttcacagagc	tgagaaactt	gaccttcctg	gacctctctc	agtgtcaact	ggagcaattg	1440
tctccaacag	catttaactc	actctccagt	cttcaggtac	taaatatgag	ccacaacaac	1500
ttcttttcat	tggatacgtt	tccttataag	tgtctgaact	ccctccaggt	tcttgattac	1560
agtctcaatc	acataatgac	ttccaaaaaa	caggaactac	agcgttttcc	aagtagtcta	1620
gccttcttaa	atcttactca	gaatgacttt	gcttgtactt	gtgaacacga	gagtttcctg	1680
cagtggatca	aggaccagag	gcagctcttg	gtggaagttg	aacgaatgga	atgtgcaaca	1740
ccttcagata	agcagggcat	gcctgtgctg	agtttgaata	tcacctgtca	gatgaataag	1800
accatcattg	gtgtgtcagt	cctcagtgtg	cttgtagtat	ctgttgtagc	agttctggtc	1860
tataagttct	attttcacct	gatgcttctt	gctggctgca	tgaagtatgg	tagaggtgaa	1920
aacacctatg	atgcctttgt	tatctactcc	agccaggatg	aggactgggt	aaggaatgag	1980
ctagtaaaga	atttagaaga	aggggtgcct	ccctttcagc	tctgccttca	ctacagagac	2040
tttattccyg	gtgtggccat	tgctgccaac	atcatccatg	aaggtttcca	taaaagccga	2100
aaggtgattg	ttgtggtgtc	ccagcacttc	atccagagcc	gctggtgtat	ctttgagtat	2160
gagattgctc	agacctggca	gtttctgagc	agtcatgctg	ggatcatctt	cattgtcctg	2220
cagaaggtgg	agaagaccct	gctcaggcag	caggtggagc	tgtaccgcct	tctcagcagg	2280
aacacttacc	tggagtggga	ggatagtgtc	ctggggcggc	acattttctg	gagacgactc	2340
agaaaagccc	tgctggatgg	taaatcatgg	aatccagaag	gaacagtggg	tacaggatgc	2400
aattag						2406

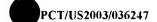


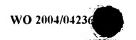


<2: <2:	10> 11> 12> 13>	8 240 DNA Hylo	ő obat∈	es la	ır											
<22	20> 21> 22>	CDS (1).	. (24	106)												
<40		8														
gtç Val 1	g gtt . Val	cct Pro	aac Asn	att Ile 5	act Thr	tat Tyr	caa Glm	tgc Cys	ato Met 10	gag Glu	r cto Lei	g aat 1 Asr	ttc Phe	tac Tyr 15	aaa Lys	48
ato	cco Pro	gac Asp	aac Asn 20	ctc Leu	Pro	ttc Phe	tca Ser	acc Thr 25	: aag " Lys	aac Asn	cto Lev	gac Asp	ctg Leu 30	agc Ser	ttt Phe	96
aat Asn	ccc Pro	ctg Leu 35	agg Arg	cat His	tta Leu	ggc Gly	agc Ser 40	tat Tyr	agc Ser	ttc Phe	ttc Phe	agt Ser 45	ttc Phe	cca Pro	gaa Glu	144
ctg Leu	Cag Gln 50	gtg Val	ctg Leu	gat Asp	tta Leu	tcc Ser 55	agg Arg	tgt Cys	gaa Glu	atc Ile	cag Gln 60	aca Thr	att	gaa Glu	gat Asp	192
ggg Gly 65	gca Ala	tat Tyr	cag Gln	agc Ser	cta Leu 70	agc Ser	ctc Leu	ctc Leu	tcc Ser	acc Thr 75	tta Leu	ata Ile	ttg Leu	aca Thr	gga Gly 80	240
aac Asn	ccc Pro	atc Ile	cag Gln	agt Ser 85	tta Leu	gct Ala	ctg Leu	gga Gly	gcc Ala 90	ttt Phe	tct Ser	gga Gly	cta Leu	tca Ser 95	agt Ser	288
tta Leu	cag Gln	aag Lys	cta Leu 100	gtg Val	gct Ala	gtg Val	gag Glu	aca Thr 105	aat Asn	cta Leu	gca Ala	tct Ser	cta Leu 110	gag Glu	aac Asn	336
ttc Phe	ccc Pro	att Ile 115	gga Gly	cat His	ctc Leu	aaa Lys	act Thr 120	ttg Leu	aaa Lys	gaa Glu	ctt Leu	aat Asn 125	gtg Val	gct Ala	cac His	384
aat Asn	ctt Leu 130	atc Ile	caa Gln	tct Ser	ttc Phe	aaa Lys 135	tta Leu	cct Pro	gag Glu	tat Tyr	ttt Phe 140	tct Ser	aat Asn	ctg Leu	acc Thr	432
aat Asn 145	cta Leu	gag Glu	cac His	ttg Leu	gac Asp 150	ctt Leu.	tcc Ser	agc Ser	aac Asn	aag Lys 155	att Ile	caa Gln	agt Ser	att Ile	tat Tyr 160	480
tgc Cys	aaa Lys	gac Asp	ttg Leu	cag Gln 165	gtt Val	cta Leu	cat His	caa Gln	atg Met 170	cc <b>c</b> Pro	cta Leu	ctc Leu	aat Asn	ctc Leu 175	tct Ser	528
tta Leu	gac Asp	ctg Leu	tcc Ser 180	ctg Leu	aac Asn	cct Pro	atg Met	aac Asn 185	ttt Phe	atc Ile	caa Gln	cca Pro	ggt Gly 190	gca Ala	ttt Phe	576

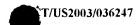


aaa Lys	a gaa s Glu	a att 11e 195	e Ser	cti Lei	crt 1 Xaa	a Lys	cto Let 200	ı Thi	tta Lei	a aga ı Arç	a aat g Asr	t aat n Asr 205	n Phe	gat Ası	agt Ser	624
tta Leu	a aat Asr 210	ı Val	ato Met	j aaa Lys	a act	tgt Cys 215	Ile	caa e Glr	ggt Gly	cto Lei	g gct 1 Ala 220	a Gly	tta Lei	gaa Glu	gtc Val	672
cat His 225	Arg	ttg Leu	gtt Val	cto Lev	g gga 1 Gly 230	/ Glu	ttt Phe	aga Arg	aat Asr	gaa Glu 235	ı Gly	a aac / Asn	tto Lev	gaa Glu	gag Glu 240	720
ttt Phe	gac Asp	: aaa Lys	Ser	gct Ala 245	Leu	gag Glu	ggc	ctg Leu	tgc Cys 250	Asn	ttç Lev	g acc n Thr	att	gaa Glu 255	gaa Glu	768
ttc Phe	cga Arg	tta Leu	gca Ala 260	Tyr	tta Leu	gac Asp	cac His	tac Tyr 265	Leu	gat Asp	gat Asp	att Ile	att Ile 270	Asp	tta Leu	816
ttt Phe	aat Asn	tgt Cys 275	ttg Leu	gca Ala	aat Asn	gtt Val	tct Ser 280	Ser	ttt Phe	tcc Ser	ctg Leu	gtg Val 285	agt Ser	gtg Val	act Thr	864
att Ile	aaa Lys 290	Arg	gta Val	gaa Glu	gac Asp	ttt Phe 295	tct Ser	tat Tyr	aat Asn	ttc Phe	gga Gly 300	Trp	caa Gln	cat His	tta Leu	912
gaa Glu 305	Leu	gtt Val	aac Asn	tgt Cys	aaa Lys 310	ttt Phe	gga Gly	cag Gln	ttt Phe	ccc Pro 315	aca Thr	ttg Leu	aac Asn	ctc Leu	aaa Lys 320	960
tct Ser	ctc Leu	aaa Lys	agg Arg	ctt Leu 325	act Thr	ttc Phe	act Thr	gcc Ala	aac Asn 330	aga Arg	ggt Gly	Gly	aat Asn	gct Ala 335	ttt Phe	1008
tca Ser	gaa Glu	gtt Val	gat Asp 340	cta Leu	cca Pro	agc Ser	ctt Leu	gag Glu 345	ttt Phe	cta Leu	gat Asp	ctc Leu	agt Ser 350	aga Arg	aat Asn	1056
ggc Gly	ttg Leu	agt Ser 355	ttc Phe	Lys	ggt Gly	Cys	Cys	Ser	caa Gln	agt Ser	gat Asp	ttt Phe 365	ggg Gly	aca Thr	aac Asn	1104
agc Ser	cta Leu 370	aag Lys	tat Tyr	tta Leu	gat Asp	ctg Leu 375	agc Ser	ttc Phe	aat 'Asn	gat Asp	gtt Val 380	att Ile	acc Thr	atg Met	agt Ser	1152
tca Ser 385	aac Asn	ttc Phe	ttg Leu	ggc Gly	tta Leu 390	gaa Glu	cag Gln	cta Leu	gaa Glu	cat His 395	ctg Leu	gat Asp	ttg Leu	cag Gln	cat His 400	1200
tcc Ser	aat Asn	ttg Leu	aaa Lys	caa Gln 405	atg Met	agt Ser	gaa Glu	ttt Phe	tca Ser 410	gta Val	ttc Phe	cta Leu	tca Ser	ctc Leu 415	aga Arg	1248
aac Asn	ctc Leu	He	tac Tyr 420	ctt Leu	gac Asp	att Ile	Ser	cat His 425	act Thr	cac His	acc Thr	Arg	gtt Val 430	gct Ala	ttc Phe	1296

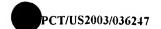




aa As	t gg n Gl	c at y Il 43	e Pr	c aane As	it gg sn Gl	c tt y Le	g tc u Se 44	r As	t ct n Le	c ga u Gl	a gt u Va	c tt l Le 44	u Ly	ia at 's Me	g gct t Ala	1344
gg Gl	y Asi 45	n se	t tt	c ca ne Gl	g ga n Gl	a aa u As: 45!	n Pho	c ct e Le	t cc u Pr	a ga o As	t at p Il 46	e Ph	c ac e Th	a ga r Gl	g ctg u Leu	1392
aga Arc 465	J ASI	c tt n Le	g ac u Th	c tt	c ct e Le 47	u Asp	c cto p Lei	c tc ı Se.	t cad	g tg n Cy: 47!	s Gl	a ct	g ga u Gl	g ca u Gl	a ttg n Leu 480	1440
tci Sei	cca Pro	a ac Th	a gc r Al	a tt a Ph 48	e Ası	c tca n Sei	a cto	tco Sei	c agi r Sei 490	r Lei	t ca u Gl:	g gta n Va	a ct l Le	a aa u As 49	t atg n Met 5	1488
ago Ser	cac His	aa As	c aa n As 50	n Ph	c tti e Phe	tca e Ser	tto Leu	g gat 1 Asp 505	Thi	g ttt c Phe	cot Pro	t tai	t aad t Ly: 510	s Cy	t ctg s Leu	1536
aac Asn	: tcc :Ser	Let 51!	u GI	g gt n Va	t ctt l Lei	gat Asp	tac Tyr 520	Ser	cto Lev	aat Asr	cad His	c ata 5 Ile 525	e Met	g act	tcc Ser	1584
aaa Lys	aaa Lys 530	GII	g gaa	a cta u Lei	a caç ı Glr	g cgt Arg 535	Phe	cca Pro	agt Ser	agt Ser	Leu 540	ı Ala	tto Phe	c tta e Leu	a aat 1 Asn	1632
ctt Leu 545	inr	caç Glr	g aat n Asi	t gad n Asp	Phe 550	Ala	tgt Cys	act Thr	tgt Cys	gaa Glu 555	His	gag Glu	agt Ser	tto Phe	ctg Leu 560	1680
cag Gln	tgg Trp	ato	aaq Lys	g gad s Asp 565	Gln	agg Arg	cag Gln	ctc Leu	ttg Leu 570	gtg Val	gaa Glu	gtt Val	gaa Glu	cga Arg 575		1728
gaa Glu	tgt Cys	gca Ala	Thr 580	Pro	tca Ser	gat Asp	aag Lys	cag Gln 585	ggc Gly	atg Met	cct Pro	gtg Val	ctg Leu 590	Ser	ttg Leu	1776
aat Asn	atc Ile	acc Thr 595	Cys	cag Gln	atg Met	aat Asn	aag Lys 600	acc Thr	atc Ile	att Ile	ggt Gly	gtg Val 605	tca Ser	gtc Val	ctc Leu	1824
agt Ser	gtg Val 610	ctt Leu	gta Val	gta Val	tct Ser	gtt Val 615	gta Val	gca Ala	gtt Val	ctg Leu	gtc Val 620	tat Tyr	aag Lys	ttc Phe	tat Tyr	1872
ttt Phe 625	cac His	ctg Leu	atg Met	ctt Leu	ctt Leu 630	gct Ala	ggc Gly	tgc Cys	atg Met	aag Lys 635	tat Tyr	ggt Gly	aga Arg	ggt Gly	gaa Glu 640	1920
aac Asn	acc Thr	tat Tyr	gat Asp	gcc Ala 645	ttt Phe	gtt Val	atc Ile	tac Tyr	tcc Ser 650	agc Ser	cag Gln	gat Asp	gag Glu	gac Asp 655	tgg Trp	1968
gta Val	agg Arg	aat Asn	gag Glu 660	cta Leu	gta Val	aag Lys	Asn	tta Leu 665	gaa Glu	gaa Glu	ggg Gly	gtg Val	cct Pro 670	ccc Pro	ttt Phe	2016



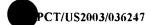
cag ctc tgc ctt cac tac aga gac ttt att ccy ggt gtg gcc att gct Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Xaa Gly Val Ala Ile Ala 675 680 685	064
gcc aac atc.atc cat gaa ggt ttc cat aaa agc cga aag gtg att gtt Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val 690 695 700	112
gtg gtg tcc cag cac ttc atc cag agc cgc tgg tgt atc ttt gag tat Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 705 710 715 720	160
gag att gct cag acc tgg cag ttt ctg agc agt cat gct ggg atc atc Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser His Ala Gly Ile Ile 725 730 735	208
ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cag cag gtg  Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val  740  745  750	256
gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gat Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 755 760 765	304
agt gtc ctg ggg cgg cac att ttc tgg aga cga ctc aga aaa gcc ctg Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 780	352
ctg gat ggt aaa tca tgg aat cca gaa gga aca gtg ggt aca gga tgc Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys 785 790 795 800	400
aat tag Asn	406
<210> 9	
<211> 801 <212> PRT <213> Hylobates lar	
<220> <221> misc_feature <222> (198)(198) <223> The 'Xaa' at location 198 stands for Arg, or His.	
<220> <221> misc_feature <222> (683)(683) <223> The 'Xaa' at location 683 stands for Pro.	
<400> 9	
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1 5 10 15	



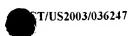
- Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 20 25 30
- Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu 35 40 45
- Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp 50 60
- Gly Ala Tyr Gln Ser Leu Ser Leu Ser Thr Leu Ile Leu Thr Gly 65 70 75 80
- Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser 85 90 95
- Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn 100 105 110
- Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His 115 120 125
- Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr 130 135 140
- Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr 145 150 155 160
- Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser 165 170 175
- Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe 180 180 190
- Lys Glu Ile Ser Leu Xaa Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser 195 200 205
- Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val 210 215 220
- His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Gly Asn Leu Glu Glu 225 235 240
- Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu 245 250 255



- Phe Arg Leu Ala Tyr Leu Asp His Tyr Leu Asp Asp Ile Ile Asp Leu 260 265 270
- Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr 275 280 285
- Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu 290 295 . 300
- Glu Leu Val Asn Cys Lys Phe Gly Gln Phe Pro Thr Leu Asn Leu Lys 305 310 315
- Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Arg Gly Gly Asn Ala Phe 325 330 335
- Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 350
- Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Asn
  355 360 365
- Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Ser 370 380
- Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Leu Gln His 385 390 395 400
- Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg 405 410 415
- Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe 420 425 430
- Asn Gly Ile Phe Asn Gly Leu Ser Asn Leu Glu Val Leu Lys Met Ala 435 440 445
- Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu 450 455 460
- Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 470 475 480
- Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met 485 490 495



- Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu 500 505 510
- Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 515 520 525
- Lys Lys Gln Glu Leu Gln Arg Phe Pro Ser Ser Leu Ala Phe Leu Asn 530 540
- Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Glu Ser Phe Leu 545 550 555 560
- Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met 565 570 575
- Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu 580 585 590
- Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu 595 600 605
- Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr 610 620
- Phe His Leu Met Leu Leu Ala Gly Cys Met Lys Tyr Gly Arg Gly Glu 625 630 635 635
- Asn Thr Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp 645 650 655
- Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe 660 665 670
- Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Xaa Gly Val Ala Ile Ala 675 680 685
- Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val 690 695 700
- Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 705 710 715 720
- Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser His Ala Gly Ile Ile 725 730 735



Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val 740 745 750

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 755 760 . 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys 785 790 795 800

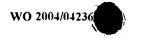
Asn

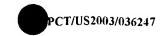
<210> 10 <211> 2388

<212> DNA

<213> Macaca mulatta

gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60 ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc 120 180 tatagettet teagttteee agaactgeag gtgetggatt tateeaggtg tgaaateeag acaattgaag atggggcata tcagagccta agccacctct ccactttaat attgacagga 240 aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 300 gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360 ttgaaagaac ttaatgtggc tcacaatctt atccagtctt tcaaattacc tgagtatttt 420 tctaatctga ccaatctaga gcacttggac ctttccagta acaagattca aaatatttat 480 tgcaaagact tgcaggttct acatcaaatg cccctatcca atctctcttt agacctgtcc 540 ctgaacccta taaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600 actttgagaa gtaattttga tgatttaaat gtaatgaaaa cttgtattca aggtctggct 660 ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaaagaaa cttggaagag 720 tttgacaaat cttctctgga gggattgtgc aatttgacca ttgaagaatt ccgattaaca 780 tacttagact actacctcga taatattatt gacttattta attgtttggc aaatgtttct 840 tcattttccc tggtgagtgt gagtattaaa agggtagaag acttttctta taatttcaga 900 tggcaacatt tagaattagt taactgtaaa tttgaacagt ttcccacatt ggaactcgaa 960





tctctcaaaa	ggcttacttt	cactgccaac	aaaggtggga	atgctttttc	agaagttgat	1020
ctaccaagco	: ttgagtttct	agatctcagt	agaaatggct	tgagtttcaa	aggttgctgt	1080
tctcaaagtg	attttgggac	aaccagccta	aagtatttag	atctgagctt	caatgatgtt	1140
attaccatga	gttcaaactt	cttgggctta	gaaaaactag	aacatctgga	tttccagcat	1200
tccaatttga	aacagatgag	tcaattttca	gtattcctat	cactcagaaa	cctcatttac	1260
cttgacattt	ctcatactca	caccagagtt	gctttcaatg	gcatcttcga	tggcttgctc	1320
agtctcaaag	tcttaaaaat	ggctggcaat	tctttccagg	aaaacttcct	tccagatatc	1380
ttcacagatc	tgaaaaactt	gaccttcctg	gacctctctc	agtgtcaatt	ggagcagttg	1440
tctccaacag	catttgacac	actcaacaag	cttcaggtac	taaatatgag	ccacaacaac	1500
ttcttttcat	tggatacgtt	tccttataag	tgtctgccct	ccctccaggt	tctcgattac	1560
agtctcaatc	acataatgac	ttccaacaac	caggaactac	agcattttcc	aagtagtcta	1620
gctttcttaa	atcttactca	gaatgacttt	gcttgtactt	gtgaacacca	gagtttcctg	1680
cagtggatca	aggaccagag	gcagctcttg	gtggaagctg	aacgaatgga	atgtgcaaca	1740
ccttcagata	aacagggcat	gccggtgctg	agtttgaata	ttacctgtca	gatgaataag	1800
accatcattg	gtgtgtctgt	gttcagtgtg	cttgtggtat	ctgttgtagc	agttctggtc	1860
tataagttct	attttcacct	gatgcttctt	gctggctgca	taaastatgg	tagaggtgaa	1920
aacatctatg	atgcctttgt	tatctactca	agccaggatg	aggactgggt	aaggaatgaa	1980
ctagtaaaga	atttagaaga	aggggtgcct	ccctttcagc	tctgccttca	ctacagagac	2040
tttattcccg	gtgtggccat	tgctgcaaac	atcatccatg	aaggtttcca	taaaagccga	2100
aaggtgattg	ttgtggtgtc	ccagcacttc	atccagagcc	gctggtgtat	ctttgaatat	2160
gagattgctc	agacctggca	gtttctgagc	agtcgtgcag	gcataatctt	cattgtcctg	2220
cagaaggtgg	agaagaccct	gctcaggcag	caggtggagc	tgtaccgcct	tctcagcagg	2280
aacacttacc	tggagtggga	ggacagtgtc	ctggggcagc	acatettetg	gagacgactc	2340
agaaaagccc	tgttggatgg	cagatcgtgg	aatccagaag	aacagtag		2388

<sup>&</sup>lt;210> 11

<sup>&</sup>lt;211> 2388

<sup>&</sup>lt;212> DNA

<sup>&</sup>lt;213> Macaca mulatta

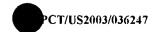
<sup>&</sup>lt;220>

<sup>&</sup>lt;221> CDS <222> (1)..(2388)



<400	)> 1	.1														
qtq	gtt	cct	aat Asn	att Ile 5	act Thr	tat Tyr	caa Gln	tgc Cys	atg Met 10	gag Glu	ctg Leu	aat Asn	ttc Phe	tac Tyr 15	aaa Lys	48
atc Ile	ccc Pro	gac Asp	aac Asn 20	ctc Leu	ccc Pro	ttc Phe	tca Ser	acc Thr 25	aag Lys	aac Asn	ctg Leu	gac Asp	ctg Leu 30	agc Ser	ttt Phe	96
aat Asn	ccc Pro	ctg Leu 35	agg Arg	cat His	tta Leu	ggc Gly	agc Ser 40	tat Tyr	agc Ser	ttc Phe	ttc Phe	agt Ser 45	ttc Phe	cca Pro	gaa Glu	144
ctg Leu	cag Gln 50	gtg Val	ctg Leu	gat Asp	tta Leu	tcc Ser 55	agg Arg	tgt Cys	gaa Glu	atc Ile	cag Gln 60	aca Thr	att Ile	gaa Glu	gat Asp	192
ggg Gly 65	gca Ala	tat Tyr	cag Gln	agc Ser	cta Leu 70	agc Ser	cac His	ctc Leu	tcc Ser	act Thr 75	tta Leu	ata Ile	ttg Leu	aca Thr	gga Gly 80	240
aac Asn	ccc Pro	atc Ile	cag Gln	agt Ser 85	tta Leu	gcc Ala	ctg Leu	gga Gly	gcc Ala 90	ttt Phe	tct Ser	gga Gly	cta Leu	tca Ser 95	agt Ser	288
tta Leu	cag Gln	aag Lys	ctg Leu 100	gtg Val	gct Ala	gtg Val	gag Glu	aca Thr 105	aat Asn	cta Leu	gca Ala	tct Ser	cta Leu 110	gag Glu	aac Asn	336
ttc Phe	ccc Pro	att Ile 115	gga Gly	cat His	ctc Leu	aaa Lys	act Thr 120	ttg Leu	aaa Lys	gaa Glu	ctt Leu	aat Asn 125	gtg Val	gct Ala	cac His	384
aat Asn	ctt Leu 130	atc Ile	cag Gln	tct Ser	ttc Phe	aaa Lys 135	tta Leu	cct Pro	gag Glu	tat Tyr	ttt Phe 140	tct Ser	aat Asn	ctg Leu	acc Thr	432
aat Asn 145	cta Leu	gag Glu	cac His	ttg Leu	gac Asp 150	ctt Leu	tcc Ser	agt Ser	aac Asn	aag Lys 155	att Ile	caa Gln	aat Asn	att Ile	tat Tyr 160	480
tgc Cys	aaa Lys	gac Asp	ttg Leu	cag Gln 165	Val	cta Leu	cat His	caa Gln	atg Met 170	ccc Pro	cta Leu	tcc Ser	aat Asn	ctc Leu 175	tct Ser	528
tta Leu	gac Asp	ctg Leu	tcc Ser 180	Leu	aac Asn	cct Pro	ata Ile	aac Asn 185	Phe	atc Ile	caa Gln	cca Pro	ggt Gly 190	gca Ala	ttt Phe	576
aaa Lys	gaa Glu	att Ile 195	Arg	ctt Leu	cat His	aag Lys	ctg Leu 200	act Thr	ttg Leu	aga Arg	agt Ser	aat Asn 205	Phe	gat Asp	gat Asp	624
tta Leu	aat Asn 210	Val	atg Met	aaa Lys	act Thr	tgt Cys 215	Ile	caa Gln	ggt Gly	ctg Leu	gct Ala 220	Gly	tta Leu	gaa Glu	gtc Val	672
cat His 225	Arg	ttg Leu	gtt Val	ctg Leu	gga Gly 230	Glu	ttt Phe	aga Arg	aat Asn	gaa Glu 235	Arg	aac Asn	ttg Leu	gaa Glu	gag Glu 240	720

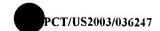




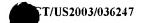
ttt Phe	gac Asp	aaa Lys	tct Ser	tct Ser 245	ctg Leu	gag Glu	gga Gly	ttg Leu	tgc Cys 250	aat Asn	ttg Leu	acc Thr	att Ile	gaa Glu 255	gaa Glu	7	768
					tta Leu											8	316
					aat Asn											8	364
att Ile	aaa Lys 290	agg Arg	gta Val	gaa Glu	gac Asp	ttt Phe 295	tct Ser	tat Tyr	aat Asn	ttc Phe	aga Arg 300	tgg Trp	caa Gln	cat His	tta Leu		912
					aaa Lys 310											<u> </u>	960
					act Thr											10	800
tca Ser	gaa Glu	gtt Val	gat Asp 340	cta Leu	cca Pro	agc Ser	ctt Leu	gag Glu 345	ttt Phe	cta Leu	gat Asp	ctc Leu	agt Ser 350	aga Arg	aat Asn	10	056
ggc Gly	ttg Leu	agt Ser 355	ttc Phe	aaa Lys	ggt Gly	tgc Cys	tgt Cys 360	tct Ser	caa Gln	agt Ser	gat Asp	ttt Phe 365	ggg	aca Thr	acc Thr	1	104
agc Ser	cta Leu 370	aag Lys	tat Tyr	tta Leu	gat Asp	ctg Leu 375	agc Ser	ttc Phe	aat Asn	gat Asp	gtt Val 380	att Ile	acc Thr	atg Met	agt Ser	1:	152
tca Ser 385	aac Asn	ttc Phe	ttg Leu	ggc Gly	tta Leu 390	gaa Glu	aaa Lys	cta Leu	gaa Glu	cat His 395	ctg Leu	gat Asp	ttc Phe	cag Gln	cat His 400	12	200
tcc Ser	aat Asn	ttg Leu	aaa Lys	cag Gln 405	atg Met	agt Ser	caa Gln	ttt Phe	tca Ser 410	gta Val	ttc Phe	cta Leu	tca Ser	ctc Leu 415	aga Arg	12	248
					gac Asp											12	296
					ggc Gly											13	344
					gaa Glu		Phe									13	392
					ctg Leu 470											14	440



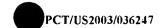
	cca Pro															1488
_	cac His						-	-	-				-	_	-	1536
	tcc Ser															1584
	aac Asn 530															1632
	act Thr															1680
	tgg Trp															1728
	tgt Cys															1776
	att Ile															1824
	gtg Val 610															1872
	cac His															1920
	atc Ile		Asp		Phe											1968
gta Val	agg Arg	aat Asn	gaa Glu 660	cta Leu	gta Val	aag Lys	aat Asn	tta Leu 665	gaa Glu	gaa Glu	ggg Gly	gtg Val	cct Pro 670	ccc Pro	ttt Phe	2016
	ctc Leu															2064
gca Ala	aac Asn 690	atc Ile	atc Ile	cat His	gaa Glu	ggt Gly 695	ttc Phe	cat His	aaa Lys	agc Ser	cga Arg 700	aag Lys	gtg Val	att Ile	gtt Val	2112
gtg Val 705	gtg Val	tcc Ser	cag Gln	cac His	ttc Phe 710	atc Ile	cag Gln	agc Ser	cgc Arg	tgg Trp 715	tgt Cys	atc Ile	ttt Phe	gaa Glu	tat Tyr 720	2160



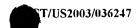
Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile 725 730 735	2208
ttc att gtc ctg cag aag gtg gag aag acc ctg ctc agg cag cag gtg Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val 740 745 750	2256
gag ctg tac cgc ctt ctc agc agg aac act tac ctg gag tgg gag gac Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 755 760 765	2304
agt gtc ctg ggg cag cac atc ttc tgg aga cga ctc aga aaa gcc ctg Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 780	2352
ttg gat ggc aga tcg tgg aat cca gaa gaa cag tag Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln 785 790 795	2388
<210> 12 <211> 795 <212> PRT <213> Macaca mulatta	
<pre>&lt;220&gt; &lt;221&gt; misc_feature &lt;222&gt; (635)(635) &lt;223&gt; The 'Xaa' at location 635 stands for Lys, or Asn.</pre>	
<400> 12	
<pre>&lt;400&gt; 12  Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1</pre>	
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys	
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1 5 10 15 Is	
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1	
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 15  Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 25  Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu 45  Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp	
Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 15  Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 25  Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu 45  Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly	



Phe	Pro	Ile 115	Gly	His	Leu	Lys	Thr 120	Leu	Lys	Glu	Leu	Asn 125	Val	Ala	His
Asn	Leu 130	Ile	Gln	Ser	Phe	Lys 135	Leu	Pro	Glu	Tyr	Phe 140	Ser	Asn	Leu	Thr
Asn 145	Leu	Glu	His	Leu	Asp 150	Leu	Ser	Ser	Asn	Lys 155	Ile	Gln	Asn	Ile	Tyr 160
Cys	Lys	Asp	Leu	Gln 165	Val	Leu	His	Gln	Met 170	Pro	Leu	Ser	Asn	Leu 175	Ser
Leu	Asp	Leu	Ser 180	Leu	Asn	Pro	Ile	Asn 185	Phe	Ile	Gln	Pro	Gly 190	Ala	Phe
Lys	Glu	Ile 195	Arg	Leu	His	Lys	Leu 200	Thr	Leu	Arg	Ser	Asn 205	Phe	Asp	Asp
Leu	Asn 210	Val	Met	Lys	Thr	Cys 215	Ile	Gln	Gly	Leu	Ala 220	Gly	Leu	Glu	Val
His 225	Arg	Leu	Val	Leu	Gly 230	Glu	Phe	Arg	Asn	Glu 235	Arg	Asn	Leu	Glu	Glu 240
Phe	Asp	Lys	Ser	Ser 245	Leu	Glu	Gly	Leu	Cys 250	Asn	Leu	Thr	Ile	Glu 255	Glu
Phe	Arg	Leu	Thr 260	Tyr	Leu	Asp	Tyr	Tyr 265	Leu	Asp	Asn	Ile	Ile 270	Asp	Leu
Phe	Asn	Cys 275	Leu	Ala	Asn	Val	Ser 280	Ser	Phe	Ser	Leu	Val 285	Ser	Val	Ser
Ile	Lys 290	Arg	Val	Glu	Asp	Phe 295	Ser	Tyr	Asn	Phe	Arg 300	Trp	Gln	His	Leu
Glu 305	Leu	Val	Asn	Cys	Lys 310	Phe	Glu	Gln	Phe	Pro 315	Thr	Leu	Glu	Leu	Glu 320
Ser	Leu	Lys	Arg	Leu 325	Thr	Phe	Thr	Ala	Asn 330	Lys	Gly	Gly	Asn	Ala 335	Phe
Ser	Glu	Val	Asp	Leu	Pro	Ser	Leu	Glu 345	Phe	Leu	Asp	Leu	Ser	Arg	Asn



Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr 355 360 Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Ser 375 Ser Asn Phe Leu Gly Leu Glu Lys Leu Glu His Leu Asp Phe Gln His 390 Ser Asn Leu Lys Gln Met Ser Gln Phe Ser Val Phe Leu Ser Leu Arg 405 410 Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe 420 425 Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala 440 435 445 Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu 455 460 450 Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 470 475 Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met 490 Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu 500 505 Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 515 520 . 525 Asn Asn Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn 530 535 Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu 560 545 550 Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met 565 570 575 Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu 580 585



Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe 595 600

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr 615 610

Phe His Leu Met Leu Leu Ala Gly Cys Ile Xaa Tyr Gly Arg Gly Glu 635 630

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala 675 680

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val 695 690

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 705 710

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val 745

Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 760 755

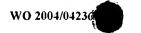
Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770

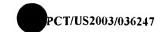
Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln 790

<210> 13 <211> 2238 <212> DNA

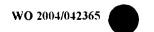
<213> Cebus capucinus

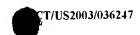
<400> 13



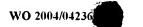


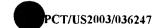
tgtgaaatcc acacaattga agatggtgca tatcagagcc taagccacct ctccacctta 60 atattgacag gaaatcctat ccagaattta gccctgqgaq ccttttctgg actatcaagt 120 ttacagaaac tggtagctgt ggagacacat ctgttatcgc tagaaagctt ccccattgga 180 catctcaaaa ctttgaagga ccttaatgtg gctcacaatc taatccaatc tttcaaatta 240 . cctgagtatt tttctaatct gaccaatcta gagcacttgg acctttctag taacaatatt 300 caaaatattt attgcaaaga cttgcaggtt ctacatcaaa tgcccctact caatctctct 360 ttagacctgt ccctgaaccc tataaacttt attcagccag gtgcatttaa agaaattagg 420 ctccgtaagc tgactttgag aaataatttt gatagtttaa atgtaatgaa aacttgcatt 480 cacggtctgg ctgggttaga agtccatcgt ttggttctgg gagaatttag aaatgaaaga 540 aatattgaag actttgacaa atctgctctg gagggcctgt gcaatttgac catcaaagaa 600 ttccgattag catacttaga caactttcca gatgatatta ttgacttatt taattgtttg 660 gtaaatgttt cttcattttc cctgttgagt gtgtatatta aaagagtaga agacttttct 720 tataatttca gatggcaaca tttagaatta gttaactgta tatttcaaca gtttcctcca 780 ctgaaactca aatctctcaa aaggcttact ttcagtaaaa acaaaggtag gaatcatttt 840 gcagaagttg atctgccaag ccttgagttt ctagatctca gtagaaatgg cttgagtttc 900 aaaggttgct gttctcaatc tgattttggg acgaccagcc taaagtattt agatctgagc 960 ttcaatgatg ttattaccat gagttcaaac ttcttaggct tagaacaact agaacacttg 1020 gatttccagc attccaattt gaaacaaatg agtgagtttt cagtatttct atcactcaga 1080 aacctcattt accttgacat ttctcatact cacaccagag ttgctttcaa tggcatcttt 1140 aatggcttgt tcagtctcaa agtcttgaaa atggctggaa attctttcca gcaaaacttc 1200 cttgcagata tcttcacaga tctgaataac ttgatattcc tggacctttc tgagtgtcaa 1260 ctggagcagt tgtctccaac agcatttgac tcacttccca gacttcagat actaaatatg 1320 agccacaaca agttetttge attggataca ttteettata agcateteta etecetecae 1380 gttctggatt acagtctcaa tcacataggg acttccaaaa atcaggaact acagcatttt 1440 ccaagtagtc tagctttctt aaatcttact caaaatgact ttgcttgtac ttgtgaacac 1500 cagagtttcc tgcagtggat caaggaccag aggcggctat tggtggaagt tgaacgaatg 1560 gaatgogoaa cacctttaaa taggaagggo atacctgtgo tgagtttgaa tatcacctgt 1620 cagatgagta agaccatcat tggtgtgtca gtgctcagtg tgcttgtggt atctgttgta 1680 gcagttctgg tctataagtt ctattttcac ctgatgcttc ttgctggctg cataaagtat 1740 ggtagaggtg aaaacaccta tgatgccttt gttatctact caagccagga tqaggactgg 1800



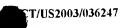


	1860
gtaaggaatg aactagtaaa gaatttagaa gaaggggtgc ctccttttca gctctgcctt	
cactacagag actttattcc cggtgtggcc attgctgcca acatcatcca tgaaggtttc	1920
cataaaagcc gaaaggtgat tgttgtggta tcccagcact tcatccagag ccgctggtgt	1980
atctttgaat atgagattge teagacetgg cagtttetga geagtegtge tggtateate	2040
ttcattgtcc tgcagaaggt ggagaagtcc ctgctcaggc agcaggtgga gctgtaccgc	2100
cttctcagca ggaacaccta cctggagtgg gaggacagtg tcctggggag gcatatcttc	2160
tggaggcgac tcagaaaagc cctgctgaat ggtagaccgt ggagtccaga aggaacagtg	2220
ggtgcaggat gcgattag	2238
<210> 14 <211> 2238	
<212> DNA	
<213> Cebus capucinus	
<220>	
<221> CDS <222> (1)(2238)	
<400> 14	4.0
tgt gaa atc cac aca att gaa gat ggt gca tat cag agc cta agc cac Cys Glu Ile His Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu Ser His	48
5 10 15	
ctc tcc acc tta ata ttg aca gga aat cct atc cag aat tta gcc ctg Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Asn Leu Ala Leu	96
20 25 30	
gga gcc ttt tct gga cta tca agt tta cag aaa ctg gta gct gtg gag Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala Val Glu	144
35 40 45	
aca cat ctg tta tcg cta gaa agc ttc ccc att gga cat ctc aaa act	192
Thr His Leu Leu Ser Leu Glu Ser Phe Pro Ile Gly His Leu Lys Thr 50 55 60	
ttg aag gac ctt aat gtg gct cac aat cta atc caa tct ttc aaa tta	240
Leu Lys Asp Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe Lys Leu 65 70 75 80	
cct gag tat ttt tct aat ctg acc aat cta gag cac ttg gac ctt tct	288
Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp Leu Ser 85 90 95	
agt aac aat att caa aat att tat tgc aaa gac ttg cag gtt cta cat	336
Ser Asn Asn Ile Gln Asn Ile Tyr Cys Lys Asp Leu Gln Val Leu His	
100	384
caa atg ccc cta ctc aat ctc tct tta gac ctg tcc ctg aac cct ata Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Ile	J0 1
115 120 125	

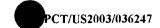




										aag Lys		432
										tgc Cys		480.
										gaa Glu 175		528
										gag Glu		576
										gac Asp		624
	_	-		-			-	-	-	gtt Val		672
										ttt Phe		720
										ttt Phe 255		768
										ttc Phe		816
										agc Ser		864
		_	-	_		_	-			 tgc Cys		912
										ctg Leu		960
										gaa Glu 335	:	1008
										agt Ser	:	1056
										att Ile	:	1104



cat His	act Thr 370	cac His	acc Thr	aga Arg	gtt Val	gct Ala 375	ttc Phe	aat Asn	ggc Gly	тте	ttt Phe 380	aat Asn	ggc Gly	ttg Leu	ttc Phe	1152
agt Ser 385	ctc Leu	aaa Lys	gtc Val	ttg Leu	aaa Lys 390	atg Met	gct Ala	gga Gly	aat Asn	tct Ser 395	ttc Phe	cag Gln	caa Gln	aac Asn	ttc Phe 400	1200
ctt Leu	gca Ala	gat Asp	atc Ile	ttc Phe 405	aca Thr	gat Asp	ctg Leu	aat Asn	aac Asn 410	ttg Leu	ata Ile	ttc Phe	ctg Leu	gac Asp 415	ctt Leu	1248
tct Ser	gag Glu	tgt Cys	caa Gln 420	ctg Leu	gag Glu	cag Gln	ttg Leu	tct Ser 425	cca Pro	aca Thr	gca Ala	ttt Phe	gac Asp 430	tca Ser	ctt Leu	1296
ccc	aga Arg	ctt Leu 435	Gln	ata Ile	cta Leu	aat Asn	atg Met 440	agc Ser	cac His	aac Asn	aag Lys	ttc Phe 445	Pne	gca Ala	ttg Leu	1344
gat Asp	aca Thr 450	Ph€	cct Pro	tat Tyr	aag Lys	cat His 455	ctc Leu	tac Tyr	tcc Ser	ctc Leu	cac His 460	var	ctg Leu	gat Asp	tac Tyr	1392
agt Ser 465	Leu	aat Asr	cac His	ata Ile	ggg Gly 470	act Thr	tcc Ser	aaa Lys	aat Asn	cag Gln 475	GIU	cta Leu	cag Gln	cat His	ttt Phe 480	1440
cca Pro	agt Ser	agt Sei	t cta r Leu	gct Ala 485	Phe	tta Leu	aat Asn	ctt Leu	act Thr 490	GIn	aat Asn	gac Asp	ttt Phe	gct Ala 495	. Cys	1488
act Thr	tgt Cys	gaa Gl	a cad u His 500	s Glr	g agt n Ser	ttc Phe	ctg Led	cag Gln 505	Trp	ato Ile	aaq Lys	g gad s Asp	caq Glr 510	HIG	g cgg g Arg	1536
cta Lev	a tto ı Lev	g gt ı Va 51	1 G1	a gtt u Val	gaa LGlu	ı cga ı Arg	ato Met 520	: GIu	tgc Cys	gca Ala	aca Thi	e cct r Pro 52!	ט בי	a aat 1 Asr	agg Arg	1584
aaq Ly:	g gg s Gl: 530	y Il	a cc e Pr	t gto o Val	g cto	g agt u Ser 535	: Le	g aat u Asr	1 116	e Thi	c Cy:	t cad s Gli	g ato n Met	g agt : Sei	aag Lys	1632
ac Th: 54	r Il	c at e Il	t gg e Gl	t gte y Va	g to: 1 Se: 55	r Val	cto Le	c agt u Sei	gto Val	g ctt L Let 55!	ı va	g gta	a tc 1 Se:	t gti r Val	t gta 1 Val 560	1680
gc Al	a gt a Va	t ct l Le	g gt u Va	c ta 1 Ty 56	r Ly	g tto s Phe	ta Ty	t tt: r Ph	t cad e Hi: 570	s Le	g at u Me	g ct t Le	t ct u Le	t gc u Al 57	t ggc a Gly 5	1728
tg Cy	c at s Il	a aa e Ly	ig ta vs Ty 58	r Gl	t ag y Ar	a ggi g Gl	t ga y Gl	a aa u As 58	n Th	c ta r Ty	t ga r As	t gc p Al	c tt a Ph 59	e va	t atc l Ile	1776
ta Ty	c to	r Se	gc ca er Gl 95	ig ga .n As	t ga p Gl	g ga u As	c tg p Tr 60	p Va	a ag l Ar	g aa g As	t ga n Gl	a ct u Le 60	u va	a aa l Ly	g aat s Asn	1824

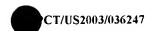


tta gaa gaa ggg Leu Glu Glu Gly 610	gtg cct cct t Val Pro Pro E 615	ttt cag ctc tgc Phe Gln Leu Cys	ctt cac tac a Leu His Tyr A 620	ga gac 1872 arg Asp
ttt att ccc ggt Phe Ile Pro Gly 625	gtg gcc att g Val Ala Ile A 630	gct gcc aac atc Ala Ala Asn Ile 635	Ile His Glu G	gt ttc 1920 Ely Phe 640
cat aaa agc cga His Lys Ser Arg	aag gtg att o Lys Val Ile V 645	gtt gtg gta tcc Val Val Val Ser 650	Gln His Phe I	tc cag 1968 le Gln 555
agc cgc tgg tgt Ser Arg Trp Cys 660				
ctg agc agt cgt Leu Ser Ser Arg 675	Ala Gly Ile 1			
aag too otg oto Lys Ser Leu Leu 690				
aac acc tac ctg Asn Thr Tyr Leu 705				
tgg agg cga ctc Trp Arg Arg Leu			Arg Pro Trp S	
gaa gga aca gtg Glu Gly Thr Val 740				2238
<210> 15 <211> 745 <212> PRT <213> Cebus cap	oucinus			
<400> 15				
Cys Glu Ile His 1	Thr Ile Glu <i>F</i> 5	Asp Gly Ala Tyr 10		er His 5
Leu Ser Thr Leu 20	Ile Leu Thr C	Gly Asn Pro Ile 25	Gln Asn Leu A	la Leu
Gly Ala Phe Ser		Ser Leu Gln Lys 40	Leu Val Ala V 45	al Glu
Thr His Leu Leu 50	Ser Leu Glu S 55	Ser Phe Pro Ile	Gly His Leu L	ys Thr



Leu	Lys	Asp	Leu	Asn	Val	Ala	His	Asn	Leu	Ile	Gln	Ser	Phe	Lys	Leu
65					70					75					80

- Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp Leu Ser 85 90 95
- Ser Asn Asn Ile Gln Asn Ile Tyr Cys Lys Asp Leu Gln Val Leu His 100 105 110
- Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Ile 115 120 125
- Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu Arg Lys Leu 130 140
- Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr Cys Ile 145 150 155 160
- His Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly Glu Phe 165 170 175
- Arg Asn Glu Arg Asn Ile Glu Asp Phe Asp Lys Ser Ala Leu Glu Gly 180 185 190
- Leu Cys Asn Leu Thr Ile Lys Glu Phe Arg Leu Ala Tyr Leu Asp Asn 195 200 205
- Phe Pro Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Val Asn Val Ser 210 215 220
- Ser Phe Ser Leu Leu Ser Val Tyr Ile Lys Arg Val Glu Asp Phe Ser 225 230 235 240
- Tyr Asn Phe Arg Trp Gln His Leu Glu Leu Val Asn Cys Ile Phe Gln 245 250 255
- Gln Phe Pro Pro Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr Phe Ser 260 265 270
- Lys Asn Lys Gly Arg Asn His Phe Ala Glu Val Asp Leu Pro Ser Leu 275 280 285
- Glu Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly Cys Cys 290 295 300



Ser Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp Leu Ser 305 310 315 320

Phe Asn Asp Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu Glu Gln 325 330 335

Leu Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met Ser Glu 340 345 350

Phe Ser Val Phe Leu Ser Leu Arg Asn Leu-Ile Tyr Leu Asp Ile Ser 355 360 365

His Thr His Thr Arg Val Ala Phe Asn Gly Ile Phe Asn Gly Leu Phe 370 375 380

Ser Leu Lys Val Leu Lys Met Ala Gly Asn Ser Phe Gln Gln Asn Phe 385 390 395 400

Leu Ala Asp Ile Phe Thr Asp Leu Asn Asn Leu Ile Phe Leu Asp Leu 405 410 415

Ser Glu Cys Gln Leu Glu Gln Leu Ser Pro Thr Ala Phe Asp Ser Leu 420 425 430

Pro Arg Leu Gln Ile Leu Asn Met Ser His Asn Lys Phe Phe Ala Leu 435 440 • 445

Asp Thr Phe Pro Tyr Lys His Leu Tyr Ser Leu His Val Leu Asp Tyr 450 455 460

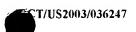
Ser Leu Asn His Ile Gly Thr Ser Lys Asn Gln Glu Leu Gln His Phe 465 470 475 480

Pro Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln Asn Asp Phe Ala Cys 485 490 495

Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln Arg Arg 500 505 510

Leu Leu Val Glu Val Glu Arg Met Glu Cys Ala Thr Pro Leu Asn Arg 515 520 525

Lys Gly Ile Pro Val Leu Ser Leu Asn Ile Thr Cys Gln Met Ser Lys 530 535 540



Thr Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser Val Val 545

Ala Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu Ala Gly 565

Cys Ile Lys Tyr Gly Arg Gly Glu Asn Thr Tyr Asp Ala Phe Val Ile 580 585 590

Tyr Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn 600

Leu Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp 615

Phe Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe 635

His Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln 650

Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe 665

Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu 675 680

Lys Ser Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg 695

Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe 710

Trp Arg Arg Leu Arg Lys Ala Leu Leu Asn Gly Arg Pro Trp Ser Pro 730

Glu Gly Thr Val Gly Ala Gly Cys Asp

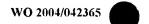
<210> 16

<211> 2406 <212> DNA

<213> Saimiri sciureus

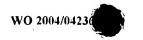
<400> 16

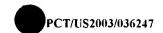
gtggttccta acgttactta	tcaatgcatg	gaactgaaty	tctacaaaat	ccccgacaac	60
atccccttct caactaagaa	cctggacctg	agctttaacc	ccctgaggca	tttaggcagc	120
catagettet teaattteee	agaactgcag	gtgctggatt	tatccaggtg	tgacatccag	180
acaatcgaag atggggcata	tcagagccta	agccacctct	ccaccttaat	attgacagga	240
aatcctatcc agaatttagc	cctgggagcc	ttttctggac	tatcaagttt	acagaagctg	300
gtggctgtgg agacacatct	gttatcacta	gagaacttcc	ccattggaca	tctcaaaact	360
ttgaaggacc ttaatgtggc	tcacaatcta	atccaatctt	tcaaattacc	tgagtatttt	420
tctaatctga ccaatctaga	gcacttggac	ctttctagta	acaatattca	aaatatttat	480
tgcaaagact tgcaggttct	acatcaaatg	cccctactca	atctctcttt	agacctgtcc	540
ctgaacccta taaactttat	tcaaccaggt	gcgtttaaag	aaattaggct	ccataagctg	600
actttgagaa ataattttga	tagtttaaat	gcaatgaaaa	cttgcattca	aggtctggct	660
gggttagaag tccatcgttt	ggttctggga	gaatttagaa	atgaaagaaa	tattgaagac	720
tttgacaaat ctgctctgga	gggcctgtgc	aatttgacca	ttaatgaatt	ccgattagct	780
tacttagatg actttctaga	tgatattatt	gacttattta	actgtttagc	aaatgtttct	840
tcattttccc tggtgaatgt	gcatattaaa	agaġtagaag	acttttctta	taattttaga	900
tggcaacatt tagaattagt	taactgtgta	tttcaacagt	ttcctccact	gaaactcaaa	960
tototoaaaa ggottacttt	cactgccaac	aaaggtagga	atcattttc	agaagttgat	1020
cttccaagcc ttgagtttct	agatctcagt	agaaatggct	tgagtttcaa	aggttgctgt	1080
tctcaatctg attttgggac	gaccagccta	aagtatttag	atctgagctt	caatgacgtt	1140
attaccatgg gttcaaactt	cttaggctta	gaacaactag	aacacttgga	tttccagcat	1200
tccaatttga aacaaatgag	tgagttttca	gtattcctat	cactcagaaa	cctcatttac	1260
cttgacattt ctcatactca	caccagagtt	gctttcaatg	gcatctttaa	tggcttgttc	1320
agtctcaaag tcttgaaaat	ggctggaaat	tctttccagc	aaaacttcct	tgaagatatc	1380
ttcacrgatc tgaataactt	gatattcctg	gacctctctg	agtgtcagct	ggagcagttg	1440
totocaacag catttgacto	acttcccaga	cttcggatac	taaatatgag	ccacaacaac	1500
ttctttgcat tggatacatt	cccttacaag	catctctact	ccctccaggt	tctggattac	1560
agtctcaatc atatagggac	ttccaaaaat	caggaactgc	agcattttcc	aagtagtcta	1620
gctttcttaa atcttactca	aaatgacttt	gcttgtactt	gtgaacacca	gagtttcctg	1680
cagtggatca aggaccagag	gcggctgttg	gtggaagttg	aacaaatgga	atgtgcaaca	1740
cctttaaata ggaagggcat	acctgtgctg	agtttgaata	tcacctgtca	gatgagtaag	1800



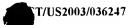


actatcattg gtgtgtcagt gctcagtgtg cttgtggtat ctgttgtagc agttctggtc	1860
tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa	1920
aacacctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgaa	1980
ctagtaaaga atttagaaga aggggtgcct ccctttcagc tctgccttca ctacagagac	2040
tttattcccg gtgtggccat tgctgccaac atcatccatg aaggtttcca taaaagccga	2100
aaggtgattg ttgtggtatc tcagcacttc atccagagcc gctggtgtat ctttgaatat	2160
gagattgete agacetggea gtttetgage agtegtgetg gtateatett cattgteetg	2220
cagaaggtgg agaagtccct getcaggcag caggtggage tgtaccgcct tetcagcagg	2280
aacacttacc tggagtggga ggacagtgtc ctggggaggc acatcttctg gagacgactc	2340
agaaaagccc tgctggatgg tagaccgtgg aatccagaag gaacagtggg tgcaggatgc	2400
gaatag	2406
<210> 17 <211> 2406 <212> DNA <213> Saimiri sciureus  <220> <221> CDS <222> (1)(2406)	
<400> 17	48
gtg gtt cct aac gtt act tat caa tgc atg gaa ctg aat ytc tac aaa Val Val Pro Asn Val Thr Tyr Gln Cys Met Glu Leu Asn Xaa Tyr Lys 1 10 15	40
atc ccc gac aac atc ccc ttc tca act aag aac ctg gac ctg agc ttt Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 20 25 30	96
aac ccc ctg agg cat tta ggc agc cat agc ttc ttc aat ttc cca gaa Asn Pro Leu Arg His Leu Gly Ser His Ser Phe Phe Asn Phe Pro Glu 35 40 45	144
ctg cag gtg ctg gat tta tcc agg tgt gac atc cag aca atc gaa gat Leu Gln Val Leu Asp Leu Ser Arg Cys Asp Ile Gln Thr Ile Glu Asp 50 55 60	192
ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly 65 70 75 80	240
aat cct atc cag aat tta gcc ctg gga gcc ttt tct gga cta tca agt	288
Asn Pro Ile Gln Asn Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser 85 90 95	

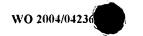


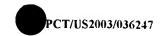


100 105 110 ttc ccc att gga cat ctc aaa act ttg aag gac ctt aat gtg gct cac 384 Phe Pro Ile Gly His Leu Lys Thr Leu Lys Asp Leu Asn Val Ala His 120 aat cta atc caa tct ttc aaa tta cct gag tat ttt tct aat ctg acc 432 Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr 135 aat cta gag cac ttg gac ctt tct agt aac aat att caa aat att tat 480 Asn Leu Glu His Leu Asp Leu Ser Ser Asn Asn Ile Gln Asn Ile Tyr 155 tgc aaa gac ttg cag gtt cta cat caa atg ccc cta ctc aat ctc tct 528 Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser 165 170 tta gac ctg tcc ctg aac cct ata aac ttt att caa cca ggt gcg ttt 576 Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe 185 aaa gaa att agg ctc cat aag ctg act ttg aga aat aat ttt gat agt 624 Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser tta aat gca atg aaa act tgc att caa ggt ctg gct ggg tta gaa gtc 672 Leu Asn Ala Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val cat cgt ttg gtt ctg gga gaa ttt aga aat gaa aga aat att gaa gac 720 His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Ile Glu Asp 230 235 ttt gac aaa tct gct ctg gag ggc ctg tgc aat ttg acc att aat gaa 768 Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Asn Glu 250 ttc cga tta gct tac tta gat gac ttt cta gat gat att att gac tta. 816 Phe Arg Leu Ala Tyr Leu Asp Asp Phe Leu Asp Asp Ile Ile Asp Leu ttt aac tgt tta gca aat gtt tct tca ttt tcc ctg gtg aat gtg cat 864 Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Asn Val His 280 att aaa aga gta gaa gac ttt tct tat aat ttt aga tgg caa cat tta 912 Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu 290 gaa tta gtt aac tgt gta ttt caa cag ttt cct cca ctg aaa ctc aaa 960 Glu Leu Val Asn Cys Val Phe Gln Gln Phe Pro Pro Leu Lys Leu Lys 315 tct ctc aaa agg ctt act ttc act gcc aac aaa ggt agg aat cat ttt 1008 Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Arg Asn His Phe tca gaa gtt gat ctt cca agc ctt gag ttt cta gat ctc agt aga aat 1056 Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn

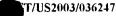


			340					345					350			
ggc Gly	ttg Leu	agt Ser 355	ttc Phe	aaa Lys	ggt. Gly	tgc Cys	tgt Cys 360	tct Ser	caa Gln	tct Ser	gat Asp	ttt Phe 365	ggg Gly	acg Thr	acc Thr	1104
agc Ser	cta Leu 370	aag Lys	tat Tyr	tta Leu	gat Asp	ctg Leu 375	agc Ser	ttc Phe	aat Asn	gac Asp	gtt Val 380	att Ile	acc Thr	atg Met	ggt Gly	1152
tca Ser 385	aac Asn	ttc Phe	tta Leu	ggc Gly	tta Leu 390	gaa Glu	caa Gln	cta Leu	gaa Glu	cac His 395	ttg Leu	gat Asp	ttc Phe	cag Gln	cat His 400	1200
tcc Ser	aat Asn	ttg Leu	aaa Lys	caa Gln 405	atg Met	agt Ser	gag Glu	ttt Phe	tca Ser 410	gta Val	ttc Phe	cta Leu	tca Ser	ctc Leu 415	aga Arg	1248
aac Asn	ctc Leu	att Ile	tac Tyr 420	ctt Leu	gac Asp	att Ile	tct Ser	cat His 425	act Thr	cac His	acc Thr	aga Arg	gtt Val 430	gct Ala	ttc Phe	1296
aat Asn	ggc Gly	atc Ile 435	ttt Phe	aat Asn	ggc Gly	ttg Leu	ttc Phe 440	agt Ser	ctc Leu	aaa Lys	gtc Val	ttg Leu 445	aaa Lys	atg Met	gct Ala	1344
gga Gly	aat Asn 450	tct Ser	ttc Phe	cag Gln	caa Gln	aac Asn 455	ttc Phe	ctt Leu	gaa Glu	gat Asp	atc Ile 460	ttc Phe	acr Xaa	gat Asp	ctg Leu	1392
aat Asn 465	aac Asn	ttg Leu	ata Ile	ttc Phe	ctg Leu 470	gac Asp	ctc Leu	tct Ser	gag Glu	tgt Cys 475	cag Gln	ctg Leu	gag Glu	cag Gln	ttg Leu 480	1440
tct Ser	cca Pro	aca Thr	gca Ala	ttt Phe 485	gac Asp	tca Ser	ctt Leu	ccc Pro	aga Arg 490	ctt Leu	cgg Arg	ata Ile	cta Leu	aat Asn 495	atg Met	1488
agc Ser	cac His	aac Asn	aac Asn 500	ttc Phe	ttt Phe	gca Ala	ttg Leu	gat Asp 505	aca Thr	ttc Phe	cct Pro	tac Tyr	aag Lys 510	cat His	ctc Leu	1536
tac Tyr	tcc Ser	ctc Leu 515	cag Gln	gtt Val	ctg Leu	gat Asp	tac Tyr 520	agt Ser	ctc Leu	aat Asn	cat His	ata Ile 525	Gly	act Thr	tcc Ser	1584
aaa Lys	aat Asn 530	cag Gln	gaa Glu	ctg Leu	cag Gln	cat His 535	ttt Phe	cca Pro	agt Ser	agt Ser	cta Leu 540	gct Ala	ttc Phe	tta Leu	aat Asn	1632
ctt Leu 545	act Thr	caa Gln	aat Asn	gac Asp	ttt Phe 550	gct Ala	tgt Cys	act Thr	tgt Cys	gaa Glu 555	His	cag Gln	agt Ser	ttc Phe	ctg Leu 560	1680
cag Gln	tgg Trp	atc Ile	aag Lys	gac Asp 565	Gln	agg Arg	cgg Arg	ctg Leu	ttg Leu 570	Val	gaa Glu	gtt Val	gaa Glu	caa Gln 575	atg Met	1728
gaa Glu	tgt Cys	gca Ala	aca Thr	cct Pro	tta Leu	aat Asn	agg Arg	aag Lys	ggc Gly	ata Ile	cct Pro	gtg Val	ctg Leu	agt Ser	ttg Leu	1776





580	0	585	590
		act atc att ggt gtg Thr Ile Ile Gly Val 605	
		gca gtt ctg gtc tat Ala Val Leu Val Tyr 620	
		tgc ata aag tat ggt Cys Ile Lys Tyr Gly 635	<del>-</del>
		tac tca agc cag gat Tyr Ser Ser Gln Asp 650	
, ,,	u Leu Val Lys Asn	tta gaa gaa ggg gtg Leu Glu Glu Gly Val 665	
		ttt att ccc ggt gtg Phe Ile Pro Gly Val 685	
		cat aaa agc cga aag His Lys Ser Arg Lys 700	
		agc cgc tgg tgt atc Ser Arg Trp Cys Ile 715	
		ctg agc agt cgt gct Leu Ser Ser Arg Ala 730	
	u Gln Lys Val Glu	aag tcc ctg ctc agg Lys Ser Leu Leu Arg 745	
gag ctg tac cgo Glu Leu Tyr Aro 755	c ctt ctc agc agg g Leu Leu Ser Arg 760	aac act tac ctg gag Asn Thr Tyr Leu Glu 765	tgg gag gac 2304 Trp Glu Asp
agt gtc ctg ggg Ser Val Leu Gly 770	g agg cac atc ttc y Arg His Ile Phe 775	tgg aga cga ctc aga Trp Arg Arg Leu Arg 780	aaa gcc ctg 2352 Lys Ala Leu
ctg gat ggt aga Leu Asp Gly Arc 785	a ccg tgg aat cca g Pro Trp Asn Pro 790	gaa gga aca gtg ggt Glu Gly Thr Val Gly 795	gca gga tgc 2400 Ala Gly Cys 800
gaa tag Glu			2406



Lys

Phe

Glu

Asp

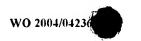
Gly 80

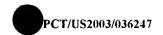
Ser

Asn

V	VO 20	104/04	2365		•									
<211 <212 <213	> P	01 RT aimi	ri s	ciur	eus									
<220 <221 <222 <223	> m	14).	feat .(14 Xaa'		loca	tion	n 14	stan	ds f	or L	eu,	or F	he.	
<220 <221 <222 <223	> m :> (	(462)	feat (4 Xaa'		loca	itior	n 462	: sta	ınds	for	Thr.			
			Ven	l eV	Thr	ጥህ r	Gln	Cvs	Met	Glu	Leu	Asn	Xaa	Tvr
1	Vai	110	AJII	5		- , -	· · ·	<b>~</b> ,-	10					15
Ile	Pro	Asp	Asn 20	Ile	Pro	Phe	Ser	Thr 25	Lys	Asn	Leu	Asp	Leu 30	Ser
Asn	Pro	Leu 35	Arg	His	Leu	Gly	Ser 40	His	Ser	Phe	Phe	Asn 45	Phe	Pro
Leu	Gln 50	Val	Leu	Asp	Leu	Ser 55	Arg	Cys	Asp	Ile	Gln 60	Thr	Ile	Glu
Gly 65	Ala	Tyr	Gln	Ser	Leu 70	Ser	His	Leu	Ser	Thr 75	Leu	Ile	Leu	Thr
Asn	Pro	Ile	Gln	Asn 85	Leu	Ala	Leu	Gly	Ala 90	Phe	Ser	Gly	Leu	Ser 95
Leu	Gln	Lys	Leu 100		Ala	Val	Glu	Thr 105	His	Leu	Leu	Ser	Leu 110	Glu

- Phe Pro Ile Gly His Leu Lys Thr Leu Lys Asp Leu Asn Val Ala His 115 120 125
- Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr 130 135 140
- Asn Leu Glu His Leu Asp Leu Ser Ser Asn Asn Ile Gln Asn Ile Tyr 145 150 150
- Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Leu Asn Leu Ser 165 170 175





Leu Asp Leu Ser Leu Asn Pro Ile Asn Phe Ile Gln Pro Gly Ala Phe 180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser 195 200 205

Leu Asn Ala Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val 210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Arg Asn Ile Glu Asp 225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Asn Glu 245 250 255

Phe Arg Leu Ala Tyr Leu Asp Asp Phe Leu Asp Asp Ile Ile Asp Leu 260 265 270

Phe Asn Cys Leu Ala Asn Val Ser Ser Phe Ser Leu Val Asn Val His 275 280 285

Ile Lys Arg Val Glu Asp Phe Ser Tyr Asn Phe Arg Trp Gln His Leu 290 295 300

Glu Leu Val Asn Cys Val Phe Gln Gln Phe Pro Pro Leu Lys Leu Lys 305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Arg Asn His Phe 325 330 335

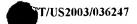
Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 350

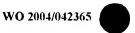
Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr 355 360 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Gly 370 380

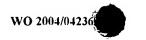
Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His 385 390 395 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg 405 410 415





- Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe
- Asn Gly Ile Phe Asn Gly Leu Phe Ser Leu Lys Val Leu Lys Met Ala
- Gly Asn Ser Phe Gln Gln Asn Phe Leu Glu Asp Ile Phe Xaa Asp Leu
- Asn Asn Leu Ile Phe Leu Asp Leu Ser Glu Cys Gln Leu Glu Gln Leu
- Ser Pro Thr Ala Phe Asp Ser Leu Pro Arg Leu Arg Ile Leu Asn Met
- Ser His Asn Asn Phe Phe Ala Leu Asp Thr Phe Pro Tyr Lys His Leu
- Tyr Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Gly Thr Ser
- Lys Asn Gln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn
- Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu
- Gln Trp Ile Lys Asp Gln Arg Arg Leu Leu Val Glu Val Glu Gln Met
- Glu Cys Ala Thr Pro Leu Asn Arg Lys Gly Ile Pro Val Leu Ser Leu
- Asn Ile Thr Cys Gln Met Ser Lys Thr Ile Ile Gly Val Ser Val Leu
- Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr
- Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu
- Asn Thr Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp



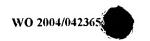


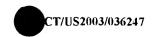
Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe 660 Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala 680 Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 715 Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu Lys Ser Leu Leu Arg Gln Gln Val 740 745 750 Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 765 755 760 Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 775 780 Leu Asp Gly Arg Pro Trp Asn Pro Glu Gly Thr Val Gly Ala Gly Cys 785 Glu <210> 19 <211> 2388 <212> DNA <213> Papio hamadryas <400> 19 gtggttccta acattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60 120 atccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc tataqcttcc tccqttttcc agaactgcag qtgctggatt tatccaggtg tgaaatccag 180 240 acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 300 aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg 360 gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact

ttgaaagaac ttaatgtggc tcacaatctt atccagtctt tcaaattacc tgagtatttt

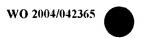


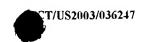
tctaatctga ccaatctaga gcacttggac ctttccagta acaagattca aaatatttat 480 tgcaaagact tgcaggttct acatcaaatg cccctaccca atctctcttt agacctgtcc 540 ctgaacccta taaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600 actttgagaa gtaattttga tgatttaaat gtaatgaaaa cttgtattca aggtctggct 660 ggtttagaag tccatcgttt ggttctggga gaatttagaa atgaaagaaa cttggaagag 720 780 tttgacaaat ctgctctgga gggattgtgc aatttgacca ttgaagaatt ccgattaaca tacttagact actacctcga taatattatt gacttattta attgtttggc aaatgcttct 840 tcattttccc tggtgagtgt gaatattaaa agggtagaag acttttctta taatttcaga 900 tggcaacatt tagaattagt taactgtaaa tttgaacagt ttcccacatt ggaactcgaa 960 tototoaaaa ggottaottt cactgocaac aaaggtggga atgootttto agaagttgat 1020 ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080 totcaaagtg attttgggac aaccagoota aagtatttag atctgagott caatgatgtt 1140 1200 attaccatgg gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat 1260 tccaatttga aacagatgag tcaattttca gtattcctat cactcagaaa cctcatttac cttgacattt ctcatactca caccacagtt gctttcaatg gcattttcga tggcttgctc 1320 agtotoaaag tottaaaaat ggotggoaat totttocagg aaaacttoot tocagatato 1380 1440 ttcacagatc tgaaaaactt gaccttcctg gacctctctc agtgtcaact ggagcagttg tctccaacag catttgacac actcaacaag cttcaggtac taaatatgag ccacaacaac 1500 ttetttteat tggatgtgtt teettataag tgtetgeeet eeetceaggt tetegattae 1560 agtotoaato acataatgao ttocaaaaao caggaacoto agcattttoo aagtagtota 1620 getttettaa atettaetea gaatgaettt gettgtaett gtgaacaeca gagttteetg 1680 cagtggatca aggaccagag gcagctcttg gtggaagctg aacgaatgga atgtgcaaca 1740 ccttcagata aacagggcat gcctgtgctg agtgtgaata ttacctgtca gatgaataag 1800 accatcattg gtgtgtctgt gttcagtgtg cttgtggtat ctgttgtagc agttctggtc 1860 tataagttct attttcacct gatgcttctt gctggctgca taaagtatgg tagaggtgaa 1920 1980 aacatctatg atgcctttgt tatctactca agccaggatg aggactgggt aaggaatgag 2040 ctagtaaaga atttagaaga aggggtgcct ccctttcagc tctgccttca ctacagagac tttattcccg gtgtggccat tgctgcaaac atcatccatg aaggtttcca taaaagccga 2100 aaggtgattg ttgtggtgtc ccagcacttc atccagagcc gctggtgtat ctttgaatat 2160 gagattgctc agacctggca gtttctgagc agtcgtgcag gcataatctt cattgtcctg 2220





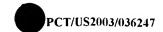
cagaaggtgg agaagaccct gctcaggcag caggtggagc tgtaccgcct tctcagcagg	2280
aacacttacc tggagtggga ggacagtgtc ctagggcagc acatcttctg gagacgactc	2340
agaaaagccc tgttggatgg cagatcgtgg aatccagaag aacagtag	2388
<210> 20 <211> 2388 <212> DNA <213> Papio hamadryas	
<220> <221> CDS <222> (1)(2388)	
<400> 20 gtg gtt cct aac att act tat caa tgc atg gag ctg aat ttc tac aaa Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1 5 10 15	48
atc ccc gac aac atc ccc ttc tca acc aag aac ctg gac ctg agc ttt Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 20 25 30	<b>96</b>
aat ccc ctg agg cat tta ggc agc tat agc ttc ctc cgt ttt cca gaa Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Leu Arg Phe Pro Glu 35 40 45	144
ctg cag gtg ctg gat tta tcc agg tgt gaa atc cag aca att gaa gat Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp 50 55 60	192
ggg gca tat cag agc cta agc cac ctc tcc acc tta ata ttg aca gga Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly 65 70 75 80	240
aac ccc atc cag agt tta gcc ctg gga gcc ttt tct gga cta tca agt Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser 85 90 95	288
tta cag aag ctg gtg gct gtg gag aca aat cta gca tct cta gag aac Leu Gln Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn 100 105 110	336
ttc ccc att gga cat ctc aaa act ttg aaa gaa ctt aat gtg gct cac Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His 115 120 125	384
aat ctt atc cag tct ttc aaa tta cct gag tat ttt tct aat ctg acc Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr 130 135 140	432
aat cta gag cac ttg gac ctt tcc agt aac aag att caa aat att tat Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Asn Ile Tyr 145 150 155 160	480
tgc aaa gac ttg cag gtt cta cat caa atg ccc cta ccc aat ctc tct Cys Lys Asp Leu Gln Val Leu His Gln Met Pro Leu Pro Asn Leu Ser	528



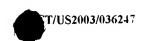


165 170 175

				165					170					1/5		
tta Leu	gac Asp	ctg Leu	tcc Ser 180	ctg Leu	aac Asn	cct Pro	ata Ile	aac Asn 185	ttt Phe	atc Ile	caa Gln	cca Pro	ggt Gly 190	gca Ala	ttt Phe	576
aaa Lys	gaa Glu	att Ile 195	agg Arg	ctt Leu	cat His	aag Lys	ctg Leu 200	act Thr	ttg Leu	aga Arg	agt Ser	aat Asn 205	ttt Phe	gat Asp	gat Asp	624
tta Leu	aat Asn 210	gta Val	atg Met	aaa Lys	act Thr	tgt Cys 215	att Ile	caa Gln	ggt Gly	ctg Leu	gct Ala 220	ggt Gly	tta Leu	gaa Glu	gtc Val	672
cat His 225	cgt Arg	ttg Leu	gtt Val	ctg Leu	gga Gly 230	gaa Glu	ttt Phe	aga Arg	aat Asn	gaa Glu 235	aga Arg	aac Asn	ttg Leu	gaa Glu	gag Glu 240	720
ttt Phe	gac Asp	aaa Lys	tct Ser	gct Ala 245	ctg Leu	gag Glu	gga Gly	ttg Leu	tgc Cys 250	aat Asn	ttg Leu	acc Thr	att Ile	gaa Glu 255	gaa Glu	768
ttc Phe	cga Arg	tta Leu	aca Thr 260	Tyr	tta Leu	gac Asp	tac Tyr	tac Tyr 265	ctc Leu	gat Asp	aat Asn	att Ile	att Ile 270	gac Asp	tta Leu	816
ttt Phe	aat Asn	tgt Cys 275	Leu	gca Ala	aat Asn	gct Ala	tct Ser 280	tca Ser	ttt Phe	tcc Ser	ctg Leu	gtg Val 285	agt Ser	gtg Val	aat Asn	864
att Ile	aaa Lys 290	Arg	gta Val	gaa Glu	gac Asp	ttt Phe 295	tct Ser	tat Tyr	aat Asn	ttc Phe	aga Arg 300	Trp	caa Gln	cat His	tta Leu	912
gaa Glu 305	Leu	gtt Val	aac Asn	tgt Cys	aaa Lys 310	Phe	gaa Glu	cag Gln	ttt Phe	ccc Pro 315	Thr	ttg Leu	gaa Glu	ctc Leu	gaa Glu 320	960
tct Ser	ctc Leu	aaa Lys	agg Arg	ctt Leu 325		ttc Phe	act Thr	gcc Ala	aac Asn 330	Lys	ggt Gly	ggg	aat Asn	gcc Ala 335	Phe	1008
tca Ser	gaa Glu	gtt Val	gat Asp 340	Let	cca Pro	ago Ser	ctt Leu	gag Glu 345	Phe	cta Leu	gat Asp	ctc Leu	agt Ser 350	Arg	aat Asn	1056
ggc	ttg Leu	agt Ser 355	: Phe	aaa Lys	a ggt s Gly	tgc Cys	tgt Cys 360	Ser	caa Gln	agt Ser	gat Asp	ttt Phe 365	e Gly	aca Thr	acc Thr	1104
ago Ser	cta Leu 370	Lys	g tat s Tyi	tta Lei	a gat u Asp	ctg Lev 375	ı Ser	tto Phe	aat Asn	gat Asp	gtt Val 380	Ile	acc Thr	ato Met	ggt Gly	1152
tca Sei 385	: Asr	tto Phe	c tto	g ggo ı Gl	c tta y Leu 390	ı Glu	ı caa ı Glr	ı cta	a gaa 1 Glu	cat His	s Let	g gat ı Asp	tto Phe	caq e Glr	g cat His 400	1200
tco Sea	aat Asr	tto Le	g aaa u Ly:	a cad	g ato n Met	g agt t Sei	caa Glr	a tti n Phe	tca e Ser	a gta : Val	a tto l Phe	c cta e Lei	a tca ı Sei	a cto Lei	aga ı Arg	1248



415 410 405 aac ctc att tac ctt gac att tct cat act cac acc aca gtt gct ttc 1296 Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Thr Val Ala Phe 425 1344 aat ggc att ttc gat ggc ttg ctc agt ctc aaa gtc tta aaa atg gct Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala 440 1392 qqc aat tot tto caq qaa aac tto ott cca gat ato tto aca gat ctg Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu 455 1440 aaa aac ttg acc ttc ctg gac ctc tct cag tgt caa ctg gag cag ttg Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 tct cca aca gca ttt gac aca ctc aac aag ctt cag gta cta aat atg 1488 Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met 490 age cae aac aac tte ttt tea ttg gat gtg ttt eet tat aag tgt etg 1536 Ser His Asn Asn Phe Phe Ser Leu Asp Val Phe Pro Tyr Lys Cys Leu 505 1584 ccc tcc ctc cag gtt ctc gat tac agt ctc aat cac ata atg act tcc Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 520 1632 aaa aac cag gaa cct cag cat ttt cca agt agt cta gct ttc tta aat Lys Asn Gln Glu Pro Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn 535 ctt act cag aat gac ttt gct tgt act tgt gaa cac cag agt ttc ctg 1680 Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu cag tgg atc aag gac cag agg cag ctc ttg gtg gaa gct gaa cga atg 1728 Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met 570 gaa tgt gca aca cct tca gat aaa cag ggc atg cct gtg ctg agt gtg 1776 Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Val 580 585 aat att acc tgt cag atg aat aag acc atc att ggt gtg tct gtg ttc 1824 Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe 600 1872 agt gtg ctt gtg gta tct gtt gta gca gtt ctg gtc tat aag ttc tat Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr 615 ttt cac ctg atg ctt ctt gct ggc tgc ata aag tat ggt aga ggt gaa 1920 Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu aac atc tat qat qcc ttt gtt atc tac tca agc cag gat gag gac tgg 1968 Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp



645 650 655

				645					650					655		
gta Val	agg Arg	aat Asn	gag Glu 660	cta Leu	gta Val	aag Lys	aat Asn	tta Leu 665	gaa Glu	gaa Glu	ggg Gly	gtg Val	cct Pro 670	ccc Pro	ttt Phe	2016
cag Gln	ctc Leu	tgc Cys 675	ctt Leu	cac His	tac Tyr	aga Arg	gac Asp 680	ttt Phe	att Ile	ccc Pro	ggt Gly	gtg Val 685	gcc Ala	att Ile	gct Ala	2064
gca Ala	aac Asn 690	atc Ile	atc Ile	cat His	gaa Glu	ggt Gly 695	ttc Phe	cat His	aaa Lys	agc Ser	cga Arg 700	aag Lys	gtg Val	att Ile	gtt Val	2112
gtg Val 705	gtg Val	tcc Ser	cag Gln	cac His	ttc Phe 710	atc Ile	cag Gln	agc Ser	cgc Arg	tgg Trp 715	tgt Cys	atc Ile	ttt Phe	gaa Glu	tat Tyr 720	2160
gag Glu	att Ile	gct Ala	cag Gln	acc Thr 725	tgg Trp	cag Gln	ttt Phe	ctg Leu	agc Ser 730	agt Ser	cgt Arg	gca Ala	ggc Gly	ata Ile 735	atc Ile	2208
ttc Phe	att Ile	gtc Val	ctg Leu 740	cag Gln	aag Lys	gtg Val	gag Glu	aag Lys 745	acc Thr	ctg Leu	ctc Leu	agg Arg	cag Gln 750	cag Gln	gtg Val	2256
gag Glu	ctg Leu	tac Tyr 755	Arg	ctt Leu	ctc Leu	agc Ser	agg Arg 760	aac Asn	act Thr	tac Tyr	ctg Leu	gag Glu 765	tgg Trp	gag Glu	gac Asp	2304
agt Ser	gtc Val 770	Leu	ggg Gly	cag Gln	cac His	atc Ile 775	ttc Phe	tgg Trp	aga Arg	cga Arg	ctc Leu 780	Arg	aaa Lys	gcc Ala	ctg Leu	2352
ttg Leu 785	Asp	ggc Gly	aga Arg	tcg Ser	tgg Trp 790	Asn	cca Pro	gaa Glu	gaa Glu	cag Gln 795						2388
c 2 1	0>	21														

<210> 21

<211> 795

<212> PRT

<213> Papio hamadryas

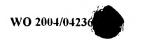
<400> 21

Val Val Pro Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys 1  $\phantom{\bigg|}$  5  $\phantom{\bigg|}$  10  $\phantom{\bigg|}$  15

Ile Pro Asp Asn Ile Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 20 25 30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Leu Arg Phe Pro Glu 35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp 50 55 60

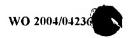


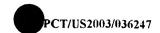


Gly 65	Ala	Tyr	Gln	Ser	Leu 70	Ser	His	Leu	Ser	Thr 75	Leu	IIe	Leu	Thr	80 GIÀ
Asn	Pro	Ile	Gln	Ser 85	Leu	Ala	Leu	Gly	Ala 90	Phe	Ser	Gly	Leu	Ser 95	Ser
Leu	Gln	Lys	Leu 100	Val	Ala	Val	Glu	Thr 105	Asn	Leu	Ala	Ser	Leu 110	Glu	Asn
Phe	Pro	Ile 115	Gly	Hiș	Leu	Lys	Thr 120	Leu	Lys	Glu	Leu	Asn 125	Val	Ala	His
Asn	Leu 130	Ile	Gln	Ser	Phe	Lys 135	Leu	Pro	Glu	Tyr	Phe 140	Ser	Asn	Leu	Thr
Asn 145	Leu	Glu	His	Leu	Asp 150	Leu	Ser	Ser	Asn	Lys 155	Ile	Gln	Asn	Ile	Туг 160
Cys	Lys	Asp	Leu	Gln 165	Val	Leu	His	Gln-	Met 170	Pro	Leu	Pro	Asn	Leu 175	Ser
Leu	Asp	Leu	Ser 180	Leu	Asn	Pro	Ile	Asn 185	Phe	Ile	Gln	Pro	Gly 190	Ala	Phe
Lys	Glu	Ile 195	Arg	Leu	His	Lys	Leu 200	Thr	Leu	Arg	Ser	Asn 205	Phe	Asp	Asp
Leu	Asn 210	Val	Меt	Lys	Thr	Cys 215	Ile	Gln	Gly	Leu ·	Ala 220	Gly	Leu	Glu	Val
His 225	Arg	Leu	Val	Leu	Gly 230	Glu	Phe	Arg	Asn	Glu 235	Arg	Asn	Leu	Glu	Glu 240
Phe	Asp	Lys ·	Ser	Ala .245	Leu	Glu	Gly	Leu	Cys 250	Asn	Leu	Thr	Ile	Glu 255	Glu
Phe	Arg	Leu	Thr 260	Tyr	Leu	Asp	Tyr	Tyr 265	Leu	Asp	Asn	Ile	Ile 270	Asp	Leu
Phe	Asn	Cys 275	Leu	Ala	Asn	Ala	Ser 280	Ser	Phe	Ser	Leu	Val 285	Ser	Val	Asn
Ile	Lys	Arg	Val	Glu	Asp	Phe 295	Ser	Tyr	Asn	Phe	Arg 300		Gln	His	Leu

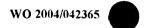


- Glu Leu Val Asn Cys Lys Phe Glu Gln Phe Pro Thr Leu Glu Leu Glu 305 310 315 320
- Ser Leu Lys Arg Leu Thr Phe Thr Ala Asn Lys Gly Gly Asn Ala Phe 325 330 335
- Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 350
- Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr 355 360 365
- Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Asp Val Ile Thr Met Gly 370 375 380
- Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His 385 390 395 400
- Ser Asn Leu Lys Gln Met Ser Gln Phe Ser Val Phe Leu Ser Leu Arg 405 410 415
- Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Thr Val Ala Phe 420 425 430
- Asn Gly Ile Phe Asp Gly Leu Leu Ser Leu Lys Val Leu Lys Met Ala 435 440 445
- Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Asp Leu 450 460
- Lys Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 470 475 480
- Ser Pro Thr Ala Phe Asp Thr Leu Asn Lys Leu Gln Val Leu Asn Met 485 490 495
- Ser His Asn Asn Phe Phe Ser Leu Asp Val Phe Pro Tyr Lys Cys Leu 500 505 510
- Pro Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 515 520 525
- Lys Asn Gln Glu Pro Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn 530 535 540





Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Ala Glu Arg Met . 570 Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Val Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Phe Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Gln His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 





Leu Asp Gly Arg Ser Trp Asn Pro Glu Glu Gln 785 790 795

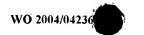
<210> 22 <211> 2427

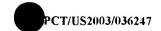
<212> DNA

<213> Pan troglodytes

<400> 22 gtggttccta atattactta tcaatgcatg gagctgaatt tctacaaaat ccccgacaac 60 120 ctccccttct caaccaagaa cctggacctg agctttaatc ccctgaggca tttaggcagc tatagettet teagttteec agaactgeag gtgetggatt tateeaggtg tgaaateeag 180 acaattgaag atggggcata tcagagccta agccacctct ccaccttaat attgacagga 240 300 aaccccatcc agagtttagc cctgggagcc ttttctggac tatcaagttt acagaagctg gtggctgtgg agacaaatct agcatctcta gagaacttcc ccattggaca tctcaaaact 360 420 ttqaaaqaac ttaatqtqqc tcacaatctt atccaatctt tcaaattacc tqaqtatttt totaatotga coaatotaga qoacttggac otttocagca acaagattoa aagtatttat 480 540 tgcacagact tgcgggttct acatcaaatg cccctactca atctctcttt agacctgtcc ctgaacccta tgaactttat ccaaccaggt gcatttaaag aaattaggct tcataagctg 600 actttgagaa ataattttga tagtttaaat gtaatgaaaa cttgtattca aggtctggct 660 qqtttagaag tccatcqttt gqttctggga gaatttagaa atgaagaaaa cttggaaaag 720 tttqacaaat ctqctctaqa qqqcctqtqc aatttqacca ttgaaqaatt ccgattagca 780 tacttaqact actacctcga tgatattatt gacttattta attgtttgac aaatgtttct 840 tcattttccc tggtgagtgt gactattaaa agcgtaaaag acttttctta taatttcgga 900 960 tqqcaacatt taqaattaqt taagtqtaaa tttggacagt ttcccacatt gaaactcaaa 1020 tototoaaaa qqottaottt cacttooaac aaaggtggga atgottttto agaagttgat ctaccaagcc ttgagtttct agatctcagt agaaatggct tgagtttcaa aggttgctgt 1080 tctcaaagtg attttgggac aaccagccta aagtatttag atctgagctt caatggtgtt 1140 attaccatga gttcaaactt cttgggctta gaacaactag aacatctgga tttccagcat 1200 tccaatttga aacaaatgag tgagttttca gtattcctat cactcagaaa cctcatttac 1260 cttgacattt ctcatactca caccagagtt gctttcaatg gcatcttcaa tggcttgtcc 1320 agtotogaag tottgaaaat ggotggcaat totttocagg aaaacttoot tocagatato 1380

ttcacaqaqc tqaqaaactt qaccttcctq qacctctctc agtgtcaact ggagcagttg



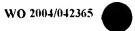


tctccaacag	cattt	taact	c ac	etete	cagt	ctt	cago	gtac	taaa	tato	gag (	ccaca	aacaac	1	500
ttcttttcat	tggat	tacgt	t to	ctta	ataaç	, tgt	ctga	act	ccct	ccaç	gt (	tctt	gattac	1	560
agtctcaatc	acata	aatga	c tt	сса	aaaa	caç	ggaad	ctac	agca	tttt	cc a	aagta	agtcta	1	620
gctttcttaa	atctt	tacto	a ga	atga	acttt	gct	tgta	actt	gtga	acac	cca a	aagtt	tcctg	1	680
caatggatca	aggad	ccaga	g go	cagct	ctto	gto	ggaag	gttg	aacq	gaato	gga a	atgt	gcaaca	1	740
ccttcagata	agcaç	gggca	t go	ctgt	gcto	g agt	ttga	ata	tcad	ctgt	ca	gatga	ataag	1	800
accatcattg	gtgt	gtcgg	t co	ctcaq	gtgtg	g ctt	gtag	gtat	ctgt	tgta	agc a	agtto	ctggtc	1	860
tataagttct	attt	tcacc	t ga	atgct	tctt	gct	ggct	gca	taaa	gtat	gg 1	tagaç	ggtgaa	1	920
aacatctatg	atgc	ctttg	t ta	atcta	actca	ago	ccago	gatg	agga	ctg	ggt a	aagga	aatgag	1	980
ctagtaaaga	attta	agaag	ra aç	ggggt	gcct	cca	attto	cagc	tctç	gcctt	ca (	ctaca	agagac	2	040
tttattcccg	gtgt	ggcca	t to	gctgo	ccaac	ațo	catco	catg	aagg	gtttc	cca	taaaa	agccga	2	100,
aaggtgattg	ttgt	ggtgt	.c c	cagca	actto	ato	ccaga	agcc	gct	gtgt	at o	cttt	gaatat	2	160
gagattgctc	agac	gtggc	a gt	tte	gago	agt	cgt	gctg	gtat	cato	ctt (	catto	gteetg	2	220 -
cagaaggtgg	agaaq	gacco	t go	ctca	ggcgg	g cad	ggtg	gagc	tgta	accgo	cct '	tctya	agcagg	2	280
aacacttacc	tgga	gtggg	ja go	gaca	gtgto	cto	3 <b>3</b> 33	egge	acat	ctto	ctg (	gagad	cgactc	2	340
agaaaagccc	tgct	ggato	g ta	aato	catgo	g aat	ccaq	gaag	gaad	agto	ggg :	tacaç	ggatgc	2	400
aattggcagg	aagca	aacat	c ta	atct	ga									2	427
<210> 23 <211> 242' <212> DNA <213> Pan	7 trog:	lodyt	es												
<220> <221> CDS <222> (1)	(242	27)													
<400> 23 gtg gtt cct Val Val Pro 1															48
atc ccc gad Ile Pro Asp	aac Asn 20	ctc Leu	ccc Pro	ttc Phe	tca Ser	acc Thr 25	aag Lys	aac Asn	ctg Leu	gac Asp	ctg Leu 30	agc Ser	ttt Phe		96
aat ccc ctc Asn Pro Lee 35	g agg ı Arg	cat His	tta Leu	ggc Gly	agc Ser 40	tat Tyr	agc Ser	ttc Phe	ttc Phe	agt Ser 45	ttc Phe	cca Pro	gaa Glu		144
ctg cag gto Leu Gln Val	g ctg L Leu	gat Asp	tta Leu	tcc Ser	agg Arg	tgt Cys	gaa Glu	atc Ile	cag Gln	aca Thr	att Ile	gaa Glu	gat Asp		192

50 55 60

	50					55					•						
ggg Gly 65	gca Ala	tat Tyr	cag Gln	agc Ser	cta Leu 70	agc Ser	cac His	ctc Leu	tcc Ser	acc Thr 75	tta Leu	ata Ile	ttg Leu	aca Thr	gga Gly 80	240	
aac Asn	ccc Pro	atc Ile	cag Gln	agt Ser 85	tta Leu	gcc Ala	ctg Leu	gga Gly	gcc Ala 90	ttt Phe	tct Ser	gga Gly	cta Leu	tca Ser 95	agt Ser	288	
tta Leu	cag Gln	aag Lys	ctg Leu 100	gtg Val	gct Ala	gtg Val	gag Glu	aca Thr 105	aat Asn	cta Leu	gca Ala	tct Ser	cta Leu 110	gag Glu	aac Asn	336	
ttc Phe	ccc Pro	att Ile 115	gga Gly	cat His	ctc Leu	aaa Lys	act Thr 120	ttg Leu	aaa Lys	gaa Glu	ctt Leu	aat Asn 125	gtg Val	gct Ala	cac His	384	i
aat Asn	ctt Leu 130	atc Ile	caa Gln	tct Ser	ttc Phe	aaa Lys 135	tta Leu	cct Pro	gag Glu	tat Tyr	ttt Phe 140	tct Ser	aat Asn	ctg Leu	acc Thr	432	?
aat Asn 145	cta Leu	gag Glu	cac His	ttg Leu	gac Asp 150	ctt Leu	tcc Ser	agc Ser	aac Asn	aag Lys 155	lle	caa Gln	agt Ser	att Ile	tat Tyr 160	480	)
tgc Cys	aca Thr	gac Asp	ttg Leu	cgg Arg 165	Val	cta Leu	cat His	caa Gln	atg Met 170	ccc Pro	cta Leu	ctc Leu	aat Asn	ctc Leu 175	Ser	528	3
tta Leu	gac Asp	ctg Leu	tcc Ser 180	Leu	aac Asn	cct Pro	atg Met	aac Asn 185	ttt Phe	atc Ile	caa Gln	cca Pro	ggt Gly 190	Ala	ttt Phe	57	6
aaa Lys	ı gaa s Glu	att Ile 195	Arg	ctt Leu	cat His	aag Lys	ctg Leu 200	Thr	ttg Leu	aga Arg	aat Asn	aat Asn 205	Phe	gat Asp	agt Ser	62	4
tta Lev	a aat u Asn 210	Va]	a atç Met	aaa Lys	act Thr	tgt Cys 215	Ile	caa Gln	ggt Gly	ctg Leu	gct Ala 220	GIA	tta Leu	gaa Glu	gtc Val	67	2
cat His 225	s Arg	tto Lev	g gtt ı Val	ctç Lev	g gga 1 Gly 230	/ Gli	ttt Phe	aga Arg	aat Asr	gaa Glu 235	GIU	a aac 1 Asr	ttç Lei	g gaa 1 Glu	a aag 1 Lys 240	72	0
tt! Ph	t gad e Asp	aaa Ly:	a tot s Sei	gct Ala 245	a Lei	a gaç ı Glu	g ggd ı Gly	c ctg / Leu	tgo Cys 250	s Asr	tto Lei	g acc	att	gaa Glu 259	a gaa ı Glu	76	8
t to Ph	c cga e Arq	tt: Le	a gca u Ala 269	а Ту	c tta r Lei	a gad ı Ası	tac Ty	tac Tyr 265	: Le	c gat ı Asp	gat Asp	t att p Ile	270	e Asi	tta Leu	81	. 6
tt Ph	t aat e Asi	t tg n Cy 27	s Le	g aca	a aat r Ası	t gti n Val	t to l Se: 28	r Sei	tti Phe	t too	c cto	g gto u Vai 285	l Se:	t gte r Val	g act l Thr	86	<b>;</b> 4
at Il	t aaa e Ly:	a ag s Se	c gt r Va	a aa l Ly	a ga s As	c tt	t tc e Se	t tat r Ty	c aar	t tte	c gg e Gl	a tg y Tr	g ca p Gl	a ca n Hi	t tta s Leu	91	.2

	290		,			295					300					
gaa Glu 305	tta Leu	gtt Val	aag Lys	tgt Cys	aaa Lys 310	ttt Phe	gga Gly	cag Gln	ttt Phe	ccc Pro 315	aca Thr	ttg Leu	aaa Lys	ctc Leu	aaa Lys 320	960
tct Ser	ctc Leu	aaa Lys	agg Arg	ctt Leu 325	act Thr	ttc Phe	act Thr	tcc Ser	aac Asn 330	aaa Lys	ggt Gly	ggg Gly	aat Asn	gct Ala 335	ttt Phe	1008
tca Ser	gaa Glu	gtt Val	gat Asp 340	cta Leu	cca Pro	agc Ser	ctt Leu	gag Glu 345	ttt Phe	cta Leu	gat Asp	ctc Leu	agt Ser 350	aga Arg	aat Asn	1056
ggc Gly	ttg Leu	agt Ser 355	ttc Phe	aaa Lys	ggt Gly	tgc Cys	tgt Cys 360	tct Ser	caa Gln	agt Ser	gat Asp	ttt Phe 365	ggg Gly	aca Thr	acc Thr	1104
agc Ser	cta Leu 370	aag Lys	tat Tyr	tta Leu	gat Asp	ctg Leu 375	agc Ser	ttc Phe	aat Asn	ggt Gly	gtt Val 380	att Ile	acc Thr	atg Met	agt Ser	1152
tca Ser 385	aac Asn	ttc Phe	ttg Leu	ggc Gly	tta Leu 390	gaa Glu	caa Gln	cta Leu	gaa Glu	cat His 395	ctg Leu	gat Asp	ttc Phe	cag Gln	cat His 400	1200
tcc Ser	aat Asn	ttg Leu	aaa Lys	caa Gln 405	atg Met	agt Ser	gag Glu	ttt Phe	tca Ser 410	gta Val	ttc Phe	cta Leu	tca Ser	ctc Leu 415	aga Arg	1248
aac Asn	ctc Leu	att Ile	tac Tyr 420	ctt Leu	gac Asp	att Ile	tct Ser	cat His 425	act Thr	cac His	acc Thr	aga Arg	gtt Val 430	gct Ala	ttc Phe	1296
aat Asn	ggc Gly	atc Ile 435	ttc Phe	aat Asn	ggc Gly	ttg Leu	tcc Ser 440	agt Ser	ctc Leu	gaa Glu	gtc Val	ttg Leu 445	aaa Lys	atg Met	gct Ala	1344
ggc Gly	aat Asn 450	tct Ser	ttc Phe	cag Gln	gaa Glu	aac Asn 455	ttc Phe	ctt Leu	cca Pro	gat Asp	atc Ile 460	ttc Phe	aca Thr	gag Glu	ctg Leu	1392
aga Arg 465	aac Asn	ttg Leu	acc Thr	ttc Phe	ctg Leu 470	gac Asp	ctc Leu	tct Ser	cag Gln	tgt Cys 475	caa Gln	ctg Leu	gag Glu	cag Gln	ttg Leu 480	1440
tct Ser	cca Pro	aca Thr	gca Ala	ttt Phe 485	Asn	tca Ser	ctc Leu	tcc Ser	agt Ser 490	ctt Leu	cag Gln	gta Val	cta Leu	aat Asn 495	atg Met	1488
agc Ser	cac His	aac Asn	aac Asn 500	ttc Phe	ttt Phe	tca Ser	ttg Leu	gat Asp 505	acg Thr	ttt Phe	cct Pro	tat Tyr	aag Lys 510	tgt Cys	ctg Leu	1536
aac Asn	tcc Ser	ctc Leu 515	Gln	gtt Val	ctt Leu	gat Asp	tac Tyr 520	agt Ser	ctc Leu	aat Asn	cac His	ata Ile 525	atg Met	act Thr	tcc Ser	1584
aaa Lys	aaa Lys	cag Gln	gaa Glu	cta Leu	cag Gln	cat His	ttt Phe	cca Pro	agt Ser	agt Ser	cta Leu	gct Ala	ttc Phe	tta Leu	aat Asn	1632



	530					535					540					
ctt Leu 545	act Thr	cag Gln	aat Asn	gac Asp	ttt Phe 550	gct Ala	tgt Cys	act Thr	tgt Cys	gaa Glu 555	cac His	caa Gln	agt Ser	ttc Phe	ctg Leu 560	1680
caa Gln	tgg Trp	atc Ile	aag Lys	gac Asp 565	cag Gln	agg Arg	cag Gln	ctc Leu	ttg Leu 570	gtg Val	gaa Glu	gtt Val	gaa Glu	cga Arg 575	atg Met	1728
gaa Glu	tgt Cys	gca Ala	aca Thr 580	cct Pro	tca Ser	gat Asp	aag Lys	cag Gln 585	ggc Gly	atg Met	cct Pro	gtg Val	ctg Leu 590	agt Ser	ttg Leu	1776
aat Asn	atc Ile	acc Thr 595	tgt Cys	cag Gln	atg Met	aat Asn	aag Lys 600	acc Thr	atc Ile	att Ile	ggt Gly	gtg Val 605	tcg Ser	gtc Val	ctc Leu	1824
agt Ser	gtg Val 610	ctt Leu	gta Val	gta Val	tct Ser	gtt Val 615	gta Val	gca Ala	gtt Val	ctg Leu	gtc Val 620	tat Tyr	aag Lys	ttc Phe	tat Tyr	1872
ttt Phe 625	cac His	ctg Leu	atg Met	ctt Leu	ctt Leu 630	gct Ala	ggc Gly	tgc Cys	ata Ile	aag Lys 635	tat Tyr	ggt Gly	aga Arg	ggt Gly	gaa Glu 640	1920
aac Asn	atc	tat Tyr	gat Asp	gcc Ala 645	Phe	gtt Val	atc Ile	tac Tyr	tca Ser 650	agc Ser	cag Gln	gat Asp	gag Glu	gac Asp 655	tgg Trp	1968
gta Val	agg Arg	aat Asn	gag Glu 660	Leu	gta Val	aag Lys	aat Asn	tta Leu 665	gaa Glu	gaa Glu	ggg Gly	gtg Val	cct Pro 670	cca Pro	ttt Phe	2016
cag Gln	g ctc Leu	tgo Cys 675	Leu	cac His	tac Tyr	aga Arg	gac Asp 680	Phe	att Ile	ccc Pro	ggt Gly	gtg Val 685	Ата	att Ile	gct Ala	2064
gcc Ala	aac Asn 690	Ile	ato lle	cat His	gaa Glu	ggt Gly 695	Phe	cat His	aaa Lys	ago	cga Arg 700	Lys	gtg Val	att	gtt Val	2112
gt ( Va ) 705	l Val	tco Ser	caç Glr	cac His	ttc Phe 710	Ile	Glr	agc Ser	cgc	tgg Trp 715	Cys	ato Ile	ttt Phe	gaa Glu	tat Tyr 720	2160
gaq Glu	g att i Ile	gct Ala	caç a Glr	acg Thr 725	Trp	caç Glr	ttt Phe	ctg Leu	ago Ser 730	Ser	cgt Arg	gct Ala	ggt Gly	ato 735	atc lle	2208
tto Pho	c att	gto Val	c cto Let 740	ı Glr	g aag n Lys	gtç Val	gaq Glu	g aag 1 Lys 745	Thi	cto Leu	g cto Leu	ago 1 Aro	g egg g Arg 750	g Gir	g gtg n Val	2256
ga Gl:	g cto u Leo	g tag 1 Ty: 75	r Ar	c ctt	cty ı Xaa	/ ago	age Are 760	g Asr	act Thi	tac Tyi	c cto	g gaq ı Glu 765	ı Trp	g gaq o Gli	g gac ı Asp	2304
ag Se	t gte r Va	c ct	g ggg u Gl	g cg y Ar	g cad	c ato	e tto	c tgg e Trp	g aga	g Ar	a cto	c aga u Arq	a aaa g Ly:	a gco s Ala	c ctg a Leu	2352

775

780

ctg gat ggt aaa tca tgg aat cca gaa gga aca gtg ggt aca gga tgc 2400 Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys 785 790 795 800

aat tgg cag gaa gca aca tct atc tga Asn Trp Gln Glu Ala Thr Ser Ilė 805 2427

<210> 24 <211> 808 <212> PRT

<213> Pan troglodytes

<220>

<221> misc\_feature <222> (758)..(758)

<223> The 'Xaa' at location 758 stands for Leu.

<400> 24

Ile Pro Asp Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe 20 25 30

Asn Pro Leu Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu 35 40 45

Leu Gln Val Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp 50 55 60

Gly Ala Tyr Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly 70 75 80

Asn Pro Ile Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser 85 90 95

Leu Gl<br/>n Lys Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn<br/>  $100 \hspace{1.5cm} 105 \hspace{1.5cm} 110 \hspace{1.5cm}$ 

Phe Pro Ile Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His 115 120 125

Asn Leu Ile Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr 130 135 140

Asn Leu Glu His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr

145 150 155 160

Cys Thr Asp Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser 165 170 175

Leu Asp Leu Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe 180 185 190

Lys Glu Ile Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser 195 200 205

Leu Asn Val Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val 210 215 220

His Arg Leu Val Leu Gly Glu Phe Arg Asn Glu Glu Asn Leu Glu Lys 225 230 235 240

Phe Asp Lys Ser Ala Leu Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu 245 250 255

Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu 260 265 270

Phe Asn Cys Leu Thr Asn Val Ser Ser Phe Ser Leu Val Ser Val Thr 275 280 285

Ile Lys Ser Val Lys Asp Phe Ser Tyr Asn Phe Gly Trp Gln His Leu 290 295 300

Glu Leu Val Lys Cys Lys Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys 305 310 315 320

Ser Leu Lys Arg Leu Thr Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe 325 330 335

Ser Glu Val Asp Leu Pro Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn 340 345 350

Gly Leu Ser Phe Lys Gly Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr 355 360 365

Ser Leu Lys Tyr Leu Asp Leu Ser Phe Asn Gly Val Ile Thr Met Ser 370 375 380

Ser Asn Phe Leu Gly Leu Glu Gln Leu Glu His Leu Asp Phe Gln His

385 390 395 400

Ser Asn Leu Lys Gln Met Ser Glu Phe Ser Val Phe Leu Ser Leu Arg 405 410 415

Asn Leu Ile Tyr Leu Asp Ile Ser His Thr His Thr Arg Val Ala Phe 420 425 430

Asn Gly Ile Phe Asn Gly Leu Ser Ser Leu Glu Val Leu Lys Met Ala 435 . 440 . 445

Gly Asn Ser Phe Gln Glu Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu 450 460

Arg Asn Leu Thr Phe Leu Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu 465 470 475 480

Ser Pro Thr Ala Phe Asn Ser Leu Ser Ser Leu Gln Val Leu Asn Met 485 490 495

Ser His Asn Asn Phe Phe Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu 500 505 510

Asn Ser Leu Gln Val Leu Asp Tyr Ser Leu Asn His Ile Met Thr Ser 515 520 525

Lys Cln Glu Leu Gln His Phe Pro Ser Ser Leu Ala Phe Leu Asn 530 540

Leu Thr Gln Asn Asp Phe Ala Cys Thr Cys Glu His Gln Ser Phe Leu 545 550 555 560

Gln Trp Ile Lys Asp Gln Arg Gln Leu Leu Val Glu Val Glu Arg Met 565 570 575

Glu Cys Ala Thr Pro Ser Asp Lys Gln Gly Met Pro Val Leu Ser Leu 580 585 590

Asn Ile Thr Cys Gln Met Asn Lys Thr Ile Ile Gly Val Ser Val Leu 595 600 605

Ser Val Leu Val Val Ser Val Val Ala Val Leu Val Tyr Lys Phe Tyr 610 620

Phe His Leu Met Leu Leu Ala Gly Cys Ile Lys Tyr Gly Arg Gly Glu

625 630 635 640

Asn Ile Tyr Asp Ala Phe Val Ile Tyr Ser Ser Gln Asp Glu Asp Trp 645 650 655

Val Arg Asn Glu Leu Val Lys Asn Leu Glu Glu Gly Val Pro Pro Phe 660 665 670

Gln Leu Cys Leu His Tyr Arg Asp Phe Ile Pro Gly Val Ala Ile Ala 675 680 685

Ala Asn Ile Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val 690 695 700

Val Val Ser Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr 705 710 715 720

Glu Ile Ala Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile 725 730 735

Phe Ile Val Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Arg Gln Val 740 745 750

Glu Leu Tyr Arg Leu Xaa Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp 755 760 765

Ser Val Leu Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu 770 775 780

Leu Asp Gly Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys 785 790 795 800

Asn Trp Gln Glu Ala Thr Ser Ile . 805

<210> 25

<211> 35

<212> PRT

<213> Amino acid

<400> 25

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr 1 5 10 15

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser 20 25 30

Phe Ser Leu 35 <210> 26 <211> 35 <212> PRT <213> Amino acid <400> 26 Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr Leu Asp Gly Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser Phe Ser Leu <210> 27 <211> 35 <212> PRT <213> Amino acid <400> 27 Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr 15 5 Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser 20 Phe Ser Leu 35 <210> 28 <211> 35 <212> PRT <213> Amino acid <400> 28 Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr

Phe Ser Leu

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cýs Leu Ala Asn Ala Ser Ser 20 25 30

<210> 29

<211> 35 <212> PRT

<213> Amino acid

<400> 29

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr 1 5

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser

Phe Ser Leu 35

<210> 30

<211> 35 <212> PRT <213> Amino acid

<400> 30

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Val Ser Ser

Phe Ser Leu 35

<210> 31

<211> 35

<212> PRT

<213> Amino acid

<400> 31

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr 10

Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser 25 20

Phe Ser Leu 35

<210> 32

<211> 35

<212> PRT

<213> Amino acid

<400> 32

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr

Leu Asp Asn Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser

Phe Ser Leu 35

<210> 33

<211> 35

<212> PRT

<213> Amino acid

<400> 33

Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Thr Tyr Leu Asp Tyr Tyr 10 15 5

Leu Asp Asn Ile Ile Asp Leu Phe Asn Cys Leu Ala Asn Ala Ser Ser 25 30

Phe Ser Leu 35

<210> 34

<211> 36

<212> PRT

<213> Amino acid

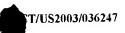
<400> 34

His Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Ile Asp Asn Tyr

Ser Ser Lys Asp Ser Ile Asp Leu Leu Asn Cys Leu Ala Asp Ile Ser 20 25

Lys Ile Ser Leu 35

<210> 35 <211> 35 <212> PRT



<213> Amino acid

<400> 35

Cys Asn Leu Thr Ile Glu Gln Phe Arg Ile Ala Tyr Leu Asp Lys Phe

Ser Gly Asp Asp Thr Asp Leu Phe Asn Cys Leu Ala Asn Val Ser Val 20 25

Ile Ser Leu 35

<210> 36 <211> 35 <212> PRT <213> Amino acid

<400> 36

Cys Asn Leu Ile Ile Glu Lys Phe Arg Ile Ala Tyr Phe Asp Lys Phe

Ser Glu Asp Ala Ile Asp Ser Phe Asn Cys Leu Ala Asn Val Ser Thr 25

Ile Ser Leu

<210> 37 <211> 35 <212> PRT <213> Amino acid

<400> 37

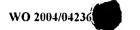
Cys Asn Leu Thr Ile Glu Lys Phe Arg Ile Ala Tyr Phe Asp Ser Phe

Ser Lys Asp Thr Thr Asn Leu Phe Asn Gln Leu Val Asn Ile Ser Ala 25

Ile Ser Leu 35

<210> 38 <211> 34 <212> PRT <213> Amino acid

<400> 38





Cys Lys Val Thr Ile Glu Glu Phe Arg Phe Thr Tyr Ala Asn Glu Phe 1 5 10 15

Ser Glu Asp Ile Thr Asp Phe Asp Cys Leu Ala Asn Val Ser Ala Met 20 25 30

Ser Leu

<210> 39

<211> 34

<212> PRT

<213> Amino acid

<400> 39

Ser Asp Asp Ile Tyr Asn Leu Asn Cys Leu Ala Asn Ile Ser Ala Met 20 25 30

Ser Phe

<210> 40

<211> 34

<212> PRT

<213> Amino acid

<400> 40

Ser Asp Asp Ile Val Lys Phe His Cys Leu Ala Asn Val Ser Ala Met 20 25 30

Ser Leu

THIS PAGE BLANK (USPTO)